Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation

April 3, 2014

M. Guardia [1] V. Kaloshin [2]

Abstract

We consider the cubic defocusing nonlinear Schrödinger equation in the two dimensional torus. Fix $s > 1$. Recently Colliander, Keel, Staffilani, Tao and Takaoka proved the existence of solutions with s-Sobolev norm growing in time.

We establish the existence of solutions with polynomial time estimates. More exactly, there is $c > 0$ such that for any $\mathcal{K} \gg 1$ we find a solution u and a time T such that $\|u(T)\|_{H^s} \geq \mathcal{K}\|u(0)\|_{H^s}$. Moreover, the time T satisfies the polynomial bound $0 < T < \mathcal{K}^c$.

Contents

[1]IAS and University of Maryland at College Park (marcel.guardia@upc.edu)
[2]IAS and University of Maryland at College Park (vadim.kaloshin@gmail.com)

1 Introduction

Let us consider the periodic cubic defocusing nonlinear Schrödinger equation (NLS),

$$\begin{cases} -i\partial_t u + \Delta u = |u|^2 u \\ u(0,x) = u_0(x) \end{cases} \tag{1.1}$$

where $x \in \mathbb{T}^2 = \mathbb{R}^2/(2\pi\mathbb{Z})^2$, $t \in \mathbb{R}$ and $u : \mathbb{R} \times \mathbb{T}^2 \to \mathbb{C}$.

The solutions of equation (1.1) conserve two quantities: the Hamiltonian

$$E[u](t) = \int_{\mathbb{T}^2} \left(\frac{1}{2} |\nabla u|^2 + \frac{1}{4} |u|^4 \right) dx(t)$$

and the mass

$$\mathcal{M}[u](t) = \int_{\mathbb{T}^2} |u|^2 dx(t) = \int_{\mathbb{T}^2} |u|^2 dx(0), \tag{1.2}$$

which is just the square of the L^2-norm of the solution, for any $t > 0$. It is useful to study the solutions $u(t)$ of (1.1) in a family of Sobolev spaces H^s with the corresponding H^s-norms

$$\|u(t)\|_{H^s(\mathbb{T}^2)} := \|u(t,\cdot)\|_{H^s(\mathbb{T}^2)} := \left(\sum_{n \in \mathbb{Z}^2} \langle n \rangle^{2s} |\hat{u}(t,n)|^2 \right)^{1/2},$$

where $\langle n \rangle = (1 + |n|^2)^{1/2}$ and,

$$\hat{u}(t,n) := \int_{\mathbb{T}^2} u(t,x) e^{-in\cdot x}\, dx.$$

The local-in-time well-posedness for any $u_0 \in H^s(\mathbb{T}^2)$, $s > 0$ was proven by Bourgain [Bou93]. This along with the two conservation laws, implies the existence of a smooth solution (1.1) for all time. It follows from the conservation of energy $E[u](t)$ that the H^1-norm of any solution of (1.1) is uniformly bounded. Our main goal is *to look for solutions whose higher Sobolev norms* $\|u(t)\|_{H^s(\mathbb{T}^2)}$, $s > 1$, *can grow in time.*

If the H^s-norm can grow indefinitely for some given $s > 1$, while the H^1-norm stays bounded, then we have solutions which initially oscillate only on scales comparable to the spatial period and eventually oscillate on arbitrarily small scales. To see that compare these norms. The only possibility for H^s to grow indefinitely is that the energy of a solution of (1.1) can penetrate to higher and higher Fourier modes.

On the one-dimensional torus, equation (1.1) is completely integrable due to the famous result of Zakharov-Shabat [ZS71] (see also [GKP12]). As a corollary $\|u(t)\|_{H^s(\mathbb{T}^1)} \leq C\|u(0)\|_{H^s(\mathbb{T}^1)}$, $s \geq 1$ for all $t > 0$. If one replaces the nonlinearity $|u|^2u = \partial_{\bar{u}}P(|u|^2)$ in (1.1) with a more general polynomial, then Bourgain [Bou96] and Staffilani [Sta97a] proved at most polynomial growth of Sobolev norms. Namely, for some $C > 0$ we have

$$\|u(t)\|_{H^s} \leq t^{C(s-1)}\|u(0)\|_{H^s} \qquad \text{for} \qquad t \to \infty.$$

In [Bou00a] Bourgain applied a version of Nekhoroshev theory. He proved that for a 1-dimensional NLS with a polynomial nonlinearity $P(|u|^2)$ satisfying $P(0) = P'(0) = P''(0) = 0$ for s large and a typical initial data $u(0) \in H^s(\mathbb{T})$ of small size ε, i.e. $\|u(0)\| \leq \varepsilon$ we have

$$\sup_{|t|<T} \|u(t)\|_{H^s} \leq C\varepsilon,$$

where $T \leq \varepsilon^{-A}$ with $A = A(s) \to 0$ as $s \to +\infty$. This is an indication of absence of a polynomial growth and motivated Bourgain [Bou00b] to pose the following question:

Are there solutions in dimension 2 *or higher with unbounded growth of* H^s*-norm for* $s > 1$*?*

Moreover, he conjectured, that in case this is true, the growth should be subpolynomial in time, that is,

$$\|u(t)\|_{H^s} \ll t^{\varepsilon}\|u(0)\|_{H^s} \qquad \text{for} \qquad t \to \infty, \text{ for all } \varepsilon > 0.$$

There are several papers obtaining improved polynomial *upper* bounds for the growth of Sobolev norms for equation (1.1) and also generalizing these results to other nonlinear Schrödinger equations either on $\mathbb{R}$, or $\mathbb{R}^2$, or on compact manifolds [Sta97b, CDKS01, Bou04, Zho08, CW10, Soh11, CKO12]. Similar results have been obtained for the wave equation [Bou96] and for the Hartree equation [Soh10b, Soh10a].

All of the cited above papers give *upper* bounds of the growth but *do not obtain* orbits which undergo growth. Indeed, there are few results obtaining such orbits. In [Bou96], Bourgain constructs orbits with unbounded growth of the Sobolev norms for the wave equation with a cubic nonlinearity but with a spectrally defined Laplacian. In [GG10, Poc11], it is shown growth of Sobolev norms for the Szegö equation, and in [Poc12] for certain nonlinear wave equation.

Concerning the nonlinear Schrödinger equation, Kuksin in [Kuk97b] (see related works [Kuk95, Kuk96, Kuk97a, Kuk99]) studied the growth of Sobolev norms but for the equation

$$-i\dot{w} = -\delta\Delta w + |w|^{2p}w, \;\; \delta \ll 1, \;\; p \geq 1.$$

He obtained solutions whose Sobolev norms grow by an inverse power of δ. Note that $u_\delta(t,x) = \delta^{-\frac{1}{2}}w(\delta^{-1}t, x)$ is a solution of (1.1). Therefore, the solutions that he obtains correspond to orbits of equation (1.1) with large initial data. The present paper is closely related to [CKS$^+$10]. In this paper, it was shown that for any $s > 1$ the H^s-norm can grow by any predetermined factor. The initial data there are not required to be large as [Kuk97b], but rather have a small initial H^s-norm with $s > 1$. *Essentially using construction from this paper* [CKS$^+$10] we not only construct solutions with similar properties, but also estimate their speed of diffusion.

The main result of this paper is

Theorem 1. *Let $s > 1$. Then there exists $c > 0$ with the following property: for any large $\mathcal{K} \gg 1$ there exists a a global solution $u(t,x)$ of (1.1) and a time T satisfying*

$$0 < T \leq \mathcal{K}^c$$

such that

$$\|u(T)\|_{H^s} \geq \mathcal{K} \|u(0)\|_{H^s}.$$

Moreover, this solution can be chosen to satisfy

$$\|u(0)\|_{L^2} \leq \mathcal{K}^{-(s-1)c/4+2/(s-1)}.$$

Note that Theorem 1 does not contradict Bourgain conjecture about the subpolynomial growth. Indeed, Theorem 1 only obtains solutions with arbitrarily large but finite growth in the Sobolev norms whereas Bourgain conjecture refers to unbounded growth.

Remark 1.1. *Even if Theorem 1 is stated for* (1.1) *in the two torus, it can be applied to the d dimensional torus with $d \geq 2$, since the solution we obtain is also a solution for equation* (1.1) *in the $\mathbb{T}^d$ setting all the other harmonics to zero.*

Remark 1.2. *In fact, we can obtain more detailed information about the distribution of the Sobolev norm of the solution $u(T)$ from Theorem 1 among its Fourier modes. More precisely, we can ensure that there exist $n_1, n_2 \in \mathbb{Z}^2$ such that*

$$\|u(T)\|_{H^s}^2 \geq |n_1|^{2s}|u_{n_1}(T)|^2 + |n_2|^{2s}|u_{n_2}(T)|^2 \geq \mathcal{K}^2\|u(0)\|_{H^s}^2.$$

That is, when $t = T$ the Sobolev norm is essentially localized on **two** *Fourier coefficients.*

Remark 1.3. *Using more careful analysis of the proof we can establish existence of solutions whose Sobolev norms are lower bounded for each time $t \in [1, T]$. Namely,*

$$\ln \|u(t)\|_{H^s} \geq \frac{t \ln \mathcal{K}}{\mathcal{K}^c} + \ln \|u(0)\|_{H^s}.$$

The solutions we construct approximate certain solutions of a finite dimensional Toy Model (see (3.12)*). The Toy Problem solutions that we use are sketched on Figure 1. Notice also that our solutions during the time interval [0,T] have two regimes:*

- *transition from one periodic solution to another one (which correspond in Figure 1 to the intersections between planes)*
- *long excursion along stable and unstable manifolds of a periodic orbit of a certain reduced system (travel through the planes).*

It turns out that during the first transition Sobolev norms grow monotonically, while during the second Sobolev norms stay practically constant.

Remark 1.4. *Our solutions differ from solutions studied in [CKS+ 10] in a substantial way. If one takes into account the information about the dynamics of the already mentioned Toy Model* (3.12) *contained in [CKS+ 10] supplied with the theory of normal forms and a beautiful trick of Shilnikov [Šil67], then it is possible to compute certain "local maps" close to some critical points and the associated diffusion time. It turns out that the diffusion time is super-exponential in $\mathcal{K}$, namely, it grows as $C^{\mathcal{K}^\alpha}$ for some $C > 0$ and $\alpha \geq 2$ (see Section 2.2 for more details).*

Even equipped with the aforementioned dynamical techniques in order to obtain polynomial diffusion time we need to achieve $\sim \ln \mathcal{K}$ cancellations in the Toy Model analyzed solutions. These cancellations are spilled out in Section 2.2 on an heuristic level and then worked out in Sections 5 and 6.

Let us just say here that the Shilnikov trick allows us to study the dynamics in a neighborhood of a certain critical point which is resonant, and therefore, not well approximated by its linearization. Thanks to this technique, we have a very precise knowledge of such dynamics, which allow us to impose these very concrete cancellations which make the growth of Sobolev norms faster.

Finally, let us point out that to achieve polynomial growth we need to ensure that the solutions of (1.1) *follow closely enough certain orbits of the Toy Model. To this end, we need to use a rather accurate approximation argument which relies on a careful choice of the modes in which the Toy Model is supported and on the precise information about the solutions of the Toy Model. This is explained in more detail in Section 2.4 and Appendix B.*

In [CKS+10] the initial conditions of solutions with growth of Sobolev norms are chosen with small $\|u(0)\|_{H^s}$[3]. In our case it is also possible, but leads to slowing down of the time of growth. This fact is explained in Appendix C (see Theorem 7).

The present paper deals with growth of Sobolev norms for a Hamiltonian partial differential equation. We show the existence of unstable solutions. As we have explained, there have not been many results showing the existence of these instabilities. In [CE10] a solution of (1.1) with spreading of mass among modes is constructed. Nevertheless the spreading does not lead to growth of Sobolev norms.

As we have already mentioned Theorem 1 is weaker than Bourgain conjecture since it asks for unbounded growth as time tends to infinity. We want to emphasize that new techniques are needed to attain unbounded growth. Indeed, the orbits we obtain are essentially supported in a finite number of modes and thus can only attain finite growth. It has been suggested that a way to obtain unbounded growth would be to concatenate solutions as those obtained in [CKS+10] and the present paper taking their support well enough separated so that, on the one hand they only weakly interact and on the other hand, the accumulation of growths leads to unbounded growth as time tends to infinity. Nevertheless, in the present paper we are only able to control the properties of such solutions for a finite time. Therefore, as time tends to infinity, such concatenated solutions may start interacting through long range convolution energy transfers regardless how far their supports are placed. Thus, as time tends to infinity, it seems rather difficult to keep track of the growth of Sobolev norms and therefore, it is not clear, how Bourgain conjecture can be proved. The only works dealing with unbounded growth are by Z. Hani [Han11, Han12]. In this papers, he shows unbounded growth for a family of pseudo partial differential equations which are a simplification of (1.1) constructed by eliminating from (1.1) precisely some long range convolution terms to overcome the problem we have just mentioned.

In the past decades there has been a considerable progress in the study of other types of dynamics for Hamiltonian partial differential equations. For instance, in the existence of periodic, quasi-periodic or almost-periodic solutions (see e.g. [Rab78, Way90, CW93, KP03, Kuk93, KP96, Ber07, BB11]), in Nekhoroshev type results (see e.g. [Bam97, Bam99]) and normal forms (see e.g. [Bam03, BG06, GIP09, GKP12, PP12]). Of particular interest for the present paper are [Bou98, EK10] since, in these papers, the authors study the existence of quasi-periodic solutions for the nonlinear Schrödinger equation in the 2-dimensional torus [Bou98] and in a torus of any dimension [EK10]. Nevertheless, they consider slightly different equations containing a convolution potential.

[3]As Terence Tao pointed to us, our solutions have small L^2-norm, but not H^s-norm

2 Main ideas and structure of the proof

One of the remarkable contributions in [CKS+10] is the formulation of a finite-dimensional Toy Model, which after a certain lift approximates some solutions of (1.1). The Hamiltonian of the Toy Model from [CKS+10] has a specific form. It has a nearest neighbors interaction and is integrable inside a certain family of 4-dimensional planes. In this section we present a class of Hamiltonians with a nearest neighbors interaction for which our method applies. It is specified at the end of Section 2.1.

2.1 Features of the model

- Write (1.1) as an infinite system of ODEs for the Fourier coefficients of the solutions. It is a Hamiltonian system with Hamiltonian $\mathcal{H}$ (see (3.2)).
- (*Two step reduction*)

  — Obtain a Normal Form of the original Hamiltonian near the origin by removing non-resonant terms (see Theorem 2).

  — Use the gauge freedom to remove the linear and some non-linear terms (see (3.7)).
- (*The Toy Model*)

  Select a finite subset of Fourier coefficients Λ in $\mathbb{Z}^2$ so that they can be split into pairwise disjoint generations $\Lambda = \cup_{j=1}^{N} \Lambda_j$ and *only neighboring* generations Λ_j and Λ_{j+1} interact.

  This can be done so that the dynamics of each element in each generation is exactly the same as the dynamics of any other member of this generation (see Corollary 3.2). Truncating we are reduced to a complex N-dimensional system given by a Hamiltonian

  $$h(b) = \frac{1}{4}\sum_{j=1}^{N} |b_j|^4 - \frac{1}{2}\sum_{j=2}^{N-1} \left(b_j^2 \bar{b}_{j-1}^2 + \bar{b}_j^2 b_{j-1}^2 \right),$$

  where each b_j is complex valued, and the symplectic form is $\Omega = \frac{i}{2} db_j \wedge \bar{b}_j$. The system conserves mass $\mathcal{M}(b) = \sum_{j=1}^{N} |b_j|^2$. We study the dynamics restricted to mass $\mathcal{M}(b) = 1$. Dynamics of this Hamiltonian is called in [CKS+10] *the Toy Model* and is the focal point of analysis. It is convenient to study this system in real coordinates and identify $\mathbb{C} \cong \mathbb{R}^2$.

  Notice also that the Hamiltonian $h(b)$ can be viewed as a Hamiltonian on a lattice $\mathbb{Z}$ with nearest neighbor interactions. Our main result relies on the construction of energy transfer from $b_3 \approx 1,\ b_j \approx 0,\ j \neq 3$ to $b_{N-1} \approx 1,\ b_j \approx 0,\ j \neq N-1$ for this Hamiltonian. Construction of a somewhat similar energy transfer for the pendulum lattice is done in [KLS11].
- (*Invariant low-dimensional subspaces*)

  Notice that each 4-dimensional plane

  $$L_j = \{b_1 = \cdots = b_{j-1} = b_{j+2} = \cdots = b_N = 0\}$$

  is invariant. Moreover, dynamics in L_j is given by a simple Hamiltonian

  $$h_j(b_j, b_{j+1}) = \frac{1}{4}\left(|b_j|^4 + |b_{j+1}|^4\right) - \frac{1}{2}\left(b_j^2 \bar{b}_{j+1}^2 + \bar{b}_j^2 b_{j+1}^2 \right).$$

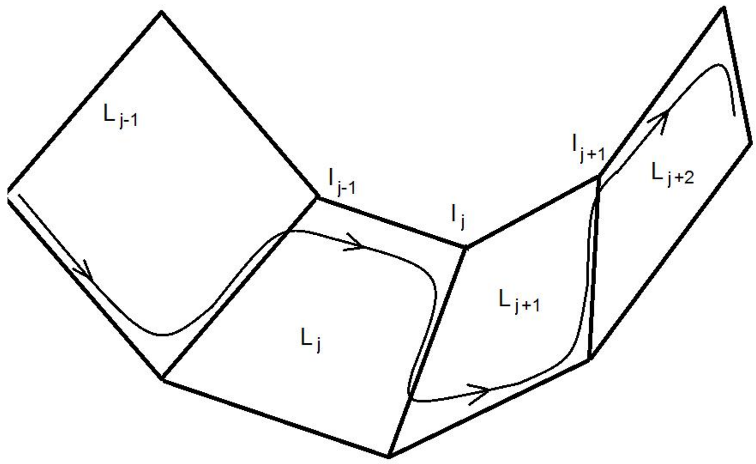

Figure 1: Planes approximating solutions

Denote $\mathcal{M}_j(b_j, b_{j+1}) = |b_j|^2 + |b_{j+1}|^2$. Both h_j and $\mathcal{M}_j$ are conserved. The mass $\mathcal{M}_j$ is assumed to be 1.

The solutions constructed stay close to the planes $\{L_j\}_{j=2}^{N-1}$ and go from a neighborhood of one intersection $l_j = L_j \cap L_{j+1}$ to a neighborhood of the next one $l_{j+1} = L_{j+1} \cap L_{j+2}$ consequently for $j = 3, \dots, N-2$ (see Figure 1).

To make a closer look at solutions we need to understand dynamics in the planes L_j's.

- (*Integrable dynamics in each plane L_j*)

  Dynamics in each 2-dimensional plane L_j is integrable. Indeed, there are two first integrals h_j and M_j in involution. By Arnold-Liouville theorem away from degeneracies the 4-dimensional plane L_j is foliated by 2-dimensional invariant tori with dynamics smoothly conjugated to a constant flow.

  We are interested in two specific periodic orbits: θ_j-direction $\{|b_j| = 1,\ b_{j+1} = 0\}$ and θ_{j+1}-direction $\{|b_{j+1}| = 1,\ b_j = 0\}$ and in a family of heteroclinic orbits $\{\gamma_j\}$ connecting the former with the later. All these orbits can be found explicitly, but *their existence can be predicted having h_j and $\mathcal{M}_j$ satisfying some properties.*

  - Having the mass $\mathcal{M}_j = |b_j|^2 + |b_{j+1}|^2$ conserved it is natural to expect that the boundary is invariant. The boundary consists of $b_j = 0$ and $b_{j+1} = 0$ (both periodic orbits), which belong to the same h_j-energy surface.
  - One can easily check that both orbits are hyperbolic, i.e. of saddle type.
  - Notice that $\{h_j = \frac{1}{4},\ \mathcal{M}_j = 1\}$ is a 2-dimensional surface with the boundary given by periodic orbits $b_j = 0$ and $b_{j+1} = 0$. Away from these periodic orbits it is a locally analytic surface, i.e. gradients ∇h_j and $\nabla \mathcal{M}_j$ are linearly independent.
  - Away from the periodic orbits $b_j = 0$ and $b_{j+1} = 0$ the surface $\{h_j = \frac{1}{4},\ \mathcal{M}_j = 1\}$ consists of stable and unstable 2-dimensional manifolds. Unless the periodic orbits

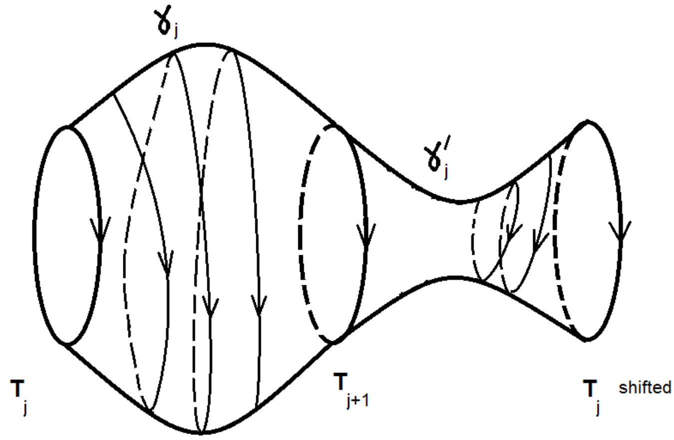

Figure 2: Heteroclinic orbits

$b_j = 0$ and $b_{j+1} = 0$ on $\{h_j = \frac{1}{4},\ \mathcal{M}_j = 1\}$ are separated by a degenerate periodic orbit, they have to be connected by these manifolds.

– Now we verify that there does not exist such a degenerate periodic orbit. Moreover, we find explicitly the family of connecting heteroclinic orbits. It turns out that these explicit formulae are not used in our proof.

Write in polar coordinates $b_k = \sqrt{r}_k\, e^{i\theta_k},\ k = j, j+1$. The mass conservation becomes $\mathcal{M}_j(b) = r_j + r_{j+1}$, the symplectic form $\Omega = \dfrac{1}{2} dr_j \wedge d\theta_j$ and the Hamiltonian

$$h_j\left(\sqrt{r_j}\; e^{i\theta_j}, \sqrt{r_{j+1}}\; e^{i\theta_{j+1}}\right) = \frac{1}{4}\left[r_j^2 + r_{j+1}^2 + 4r_j r_{j+1}\cos 2(\theta_j - \theta_{j+1})\right].$$

Then the equation of motion are

$$\begin{aligned}
\dot\theta_j &= r_j - 2r_{j+1}\cos 2(\theta_j - \theta_{j+1})\\
\dot\theta_{j+1} &= r_{j+1} - 2r_j\cos 2(\theta_j - \theta_{j+1})\\
\dot r_j &= 4r_j\, r_{j+1}\sin 2(\theta_j - \theta_{j+1})\\
\dot r_{j+1} &= -4r_j\, r_{j+1}\sin 2(\theta_j - \theta_{j+1}).
\end{aligned}$$

For the energy surface $h_j = \dfrac{1}{4}$ we have

– Two families of periodic solutions $\{(\theta_j, \theta_{j+1}, r_j, r_{j+1}):\ r_j = 0\}$ and $\{(\theta_j, \theta_{j+1}, r_j, r_{j+1}):\ r_{j+1} = 0\}$.

– Each family has two special solutions: $2(\theta_j - \theta_{j+1})$ equals either $\frac{2\pi}{3}$ and $\frac{4\pi}{3}$. Both planes are invariant: $\frac{d}{dt}(\theta_j - \theta_{j+1}) = -(r_j + r_{j+1})(1 + 2\cos 2(\theta_j - \theta_{j+1})) = 0$. Denote $\mathbb{T}_j = \{2(\theta_j - \theta_{j+1}) = \frac{2\pi}{3}\ (\bmod 2\pi),\ r_j = 0\}$.

– On $\mathcal{M}_j = 1,\ h_j = \frac{1}{4},\ \theta_j - \theta_{j+1} = \frac{2\pi}{3}$ we have $\dot r_j = r_j r_{j+1} = -\dot r_{j+1}$. Thus, there is a heteroclinic orbit γ_j connecting $\mathbb{T}_j$ with the second family $r_{j+1} = 0$.

Now we can be more specific in the location of the orbits:

$$\begin{array}{c}\textit{The solutions constructed go from one periodic orbit } \mathbb{T}_2 \textit{ to the next } \mathbb{T}_3 \\ \textit{along } \gamma_2, \textit{ then from } \mathbb{T}_3 \textit{ to } \mathbb{T}_4 \textit{ along } \gamma_3 \textit{ and so on for } j = 4, \dots, N-2.\end{array} \tag{2.1}$$

In a view of the above discussion we have the following description:

$$\begin{array}{llll} \rightsquigarrow & \rightsquigarrow\ \mathbb{T}_j\ \rightsquigarrow & \rightsquigarrow\ \gamma_j\ \rightsquigarrow & \rightsquigarrow\ \mathbb{T}_{j+1}\quad \rightsquigarrow \\ & \dot\theta_i \approx 0,\ |i-j|>1 & \theta_j - \theta_{j+1} \approx \dfrac{\pi}{3} & \dot\theta_i \approx 0,\ |i-j-1|>1 \\ & |b_i| \approx 0,\ i \neq j, j+1 & |b_j| \rightsquigarrow |b_{j+1}| & |b_i| \approx 0,\ i \neq j+1, j+2. \end{array} \tag{2.2}$$

- (*Local behavior of periodic orbits* $\mathbb{T}_j$) Due to the above analysis, the periodic orbits $\mathbb{T}_j$ viewed in $\mathbb{R}^{2N}$ have at least two expanding and two contracting directions: one pair from L_{j-1}-plane and the other from L_j-pane. Due to symmetry of the restricted systems in L_{j-1}-plane and L_j-plane these periodic orbits have **multiple** hyperbolic eigenvalues. Multiplicity turns out to be **exactly** 2.
- (*Resonant normal forms near* $\mathbb{T}_j$) The presence of the resonance complicates the analysis of the local map since, as formulae (4.37) show, the resonance modifies the local behavior compared to the linear case. To overcome this problem, we use a beautiful trick of Shilnikov [Šil67] and obtain precise information about the local behavior, which is explained in Section 2.2.
- (*Connecting heteroclinic orbits*) As we have showed above, there are orbits γ_j connecting $\mathbb{T}_j$ with $\mathbb{T}_{j+1}$ for each $j = 3, \dots, n-2$. We need to analyze the dynamics near these heteroclinic orbits.
- (*Local almost product structure*) Once we obtain information about the behavior near $\mathbb{T}_j$'s and near the connecting orbits γ_j, we can describe the dynamics of the Toy Model as if it close to the direct product of $(N-3)$ planes $L_j,\ j = 3, \dots, N-1$.

Properties of the Hamiltonian $h(b)$ used in the proof.

As we mentioned in the introduction to this section we do not use the specific form of h. Here is the list of properties that we need.

- h has nearest neighbors interaction;
- h has 2-dimensional (complex) invariant planes intersecting transversally;
- there are two first integrals (coming from two conserved quantities: energy and mass);
- some generic properties of h and $\mathcal{M}$.

Growth of Sobolev norms through resonant structures, as happens for the cubic defocusing nonlinear Schrödinger equation, is expected to take place in a large set of Hamiltonian Partial Differential Equations. For instance, in the nonlinear wave equation, the nonlinear quantum harmonic oscillator or the Hartree equation. It is not clear for the authors how the I-team approach can be implemented in such equations to obtain a Toy Model similar to the one considered in [CKS+10] and in the present paper. Nevertheless, we want to emphasize that, if a Toy Model for such equations could be obtaine, one would not need to have a very precise knowledge of its dynamics but it would suffice that it satisfies the just listed properties.

2.2 The dynamics close to the periodic orbits: a heuristic model

One of the crucial steps in analyzing the Toy Model $h(b)$ is the study of the dynamics in a neighborhood of the periodic orbits $\mathbb{T}_j$. Namely, we want to analyze how points which lie close their stable invariant manifold evolve under the flow until reaching points close to their unstable one (see Figure 3). As we have explained, these periodic orbits are of mixed type (four eigenvalues are hyperbolic and the rest are elliptic). Since in each plane L_j dynamics is the same explained in the previous section, the hyperbolic eigenvalues have multiplicity two and, therefore, are equal to $\lambda, \lambda, -\lambda, -\lambda$ for some $\lambda > 0$. Since this section serves an expository purpose, we let $\lambda = 1$ and set the elliptic modes to zero. [4]

Essentially the study has three steps:

- Using conservation of $\mathcal{M}$, make a simplectic reduction so the periodic orbit $\mathbb{T}_j$ becomes a fixed point.
- Perform a normal form procedure to reduce the size of the higher order non-resonant terms.
- Analyze the dynamics of the new vector field and achieve a cancelation for a local map.

The first step is performed in Section 4.1. It leads to a Hamiltonian of two degrees of freedom of the form

$$H(p,q) = p_1 q_1 + p_2 q_2 + H_4(q,p),$$

where H_4 is a homogeneous polynomial of degree four. The variables (p_1, q_1) correspond to the variable b_{j-1} after diagonalizing the saddle and the variables (q_2, p_2) correspond to b_{j+1}.

Fix a small $\sigma > 0$. To study the local dynamics, it suffices to analyze a map from a section $\Sigma_- = \{q_1 = \sigma,\ |p_1|, |q_2|, |p_2| \ll \sigma\}$, to a section $\Sigma_+ = \{p_2 = \sigma,\ |p_1|, |q_1|, |q_2| \ll \sigma\}$ (see Figure 3). Using rescaling assume $\sigma = 1$. This only changes time by a fixed factor.

Since we are in a neighborhood of the origin, one would expect that the dynamics of the system associated to this Hamiltonian is well approximated by its first order, that is, by a linear equation. Then, the solutions are just given by

$$\begin{aligned} p_1(t) &= p_1^0 e^t, \quad q_1(t) = q_1^0 e^{-t} \\ p_2(t) &= p_2^0 e^t, \quad q_2(t) = q_2^0 e^{-t} \end{aligned}$$

and then the local map $\mathcal{B}_0$ from $U \subset \Sigma_-$ to Σ_+ for this system sends points

$$\left(p_1^0, q_1^0, p_2^0, q_2^0\right) \sim \left(\delta, 1, \sqrt{\delta}, \sqrt{\delta}\right)$$

to

$$\mathcal{B}_0 \left(p_1^0, q_1^0, p_2^0, q_2^0\right) \sim \left(\sqrt{\delta}, \sqrt{\delta}, 1, \delta\right),$$

where $0 < \delta \ll 1$. Moreover, the travel time of orbits by this map is always $T = -\ln \sqrt{\delta} + \mathcal{O}(1)$.

[4] To be more precise near each saddle, the elliptic directions remain almost constant and, since they will be taken small enough, it turns out they do not make much influence in the dynamics of hyperbolic components. Thus, to simplify the exposition, we set the elliptic modes to zero and study how the hyperbolic ones evolve. This implies that we only need to study three modes b_{j-1}, b_j and b_{j+1}. This analysis is performed in Section 5 in great detail.

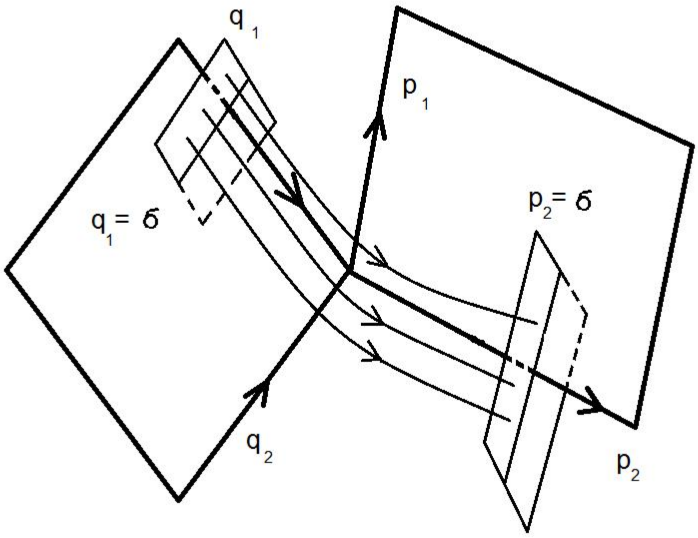

Figure 3: Local map

We will see that the image point *changes substantially* when we add H_4 to the system, due to *both resonant and nonresonant terms*. To exemplify this, we consider *a simplified model* which in fact contains *all the difficulties* that the true model has,

$$H(p,q) = p_1q_1 + p_2q_2 + q_1^2p_2^2 + p_1^2p_2^2. \tag{2.3}$$

Since the term $p_1^2p_2^2$ is nonresonant, we first perform one step of normal form $(x,y) = \Psi(p,q)$ (see Section 5 for details). It can be easily seen that the change Ψ is of the form

$$\Psi(p,q) = \left(p_1, q_1 + \mathcal{O}(p_1p_2^2), p_2, q_2 + \mathcal{O}(p_1^2p_2)\right) \tag{2.4}$$

and, therefore, keeps the size of initial points of the form

$$\left(p_1^0, q_1^0, p_2^0, q_2^0\right) \sim \left(\delta, 1, \sqrt{\delta}, \sqrt{\delta}\right).$$

That is, $(x^0, y^0) = \Psi(p^0, q^0)$ satisfies

$$\left(x_1^0, y_1^0, x_2^0, y_2^0\right) \sim \left(\delta, 1, \sqrt{\delta}, \sqrt{\delta}\right).$$

The change to normal form leads to a Hamiltonian system of the form

$$H'(x,y) = x_1y_1 + x_2y_2 + y_1^2x_2^2 + \text{higher order terms.}$$

Drop the higher order terms. Then, the solutions of the system associated to this Hamiltonian can be computed explicitly and are given by

$$\begin{aligned}
x_1 &= x_1^0e^t + 2y_1^0(x_2^0)^2te^t = \left(x_1^0 + 2y_1^0(x_2^0)^2t\right)e^t \\
y_1 &= y_1^0e^{-t} \\
x_2 &= x_2^0e^t \\
y_2 &= y_2^0e^{-t} - 2(y_1^0)^2x_2^0te^t.
\end{aligned}$$

Thus, since the travel time is $t = -\ln \sqrt{\delta} + \mathcal{O}(1)$, it is clear that *the nonlinear terms are bigger than the linear ones*, leading to an image point of the form

$$\left(x_1^f, y_1^f, x_2^f, y_2^f\right) \sim \left(\sqrt{\delta}\ln(1/\delta), \sqrt{\delta}, 1, \delta\ln(1/\delta)\right).$$

Using (2.4), in the original variables the image point of the map $\mathcal{B}_1$ associated to Hamiltonian H is of the form

$$\mathcal{B}_1\left(p_1^0, q_1^0, p_2^0, q_2^0\right) \sim \left(\sqrt{\delta}\ln(1/\delta), \sqrt{\delta}, 1, \delta\ln^2(1/\delta)\right).$$

We want to emphasize that the presence of these *logarithmic terms is a serious problem* we need to deal with. Recall that we need to travel through $N-3$ saddles ($\mathbb{T}_3 \rightsquigarrow \mathbb{T}_4 \rightsquigarrow \ldots \rightsquigarrow \mathbb{T}_{N-1}$). Roughly speaking, this implies that we need to compose $N-4$ local maps. Thanks to the symmetries, at each saddle we can consider a system of coordinates such that the dynamics is essentially given by a Hamiltonian of the form (2.3). Moreover, since at each local map we gain some logarithms, the initial points of the local map associated to the j saddle are of the form

$$\left(p_1^0, q_1^0, p_2^0, q_2^0\right) \sim \left(\delta\ln^{2^{j-1}}(1/\delta), 1, \sqrt{\delta}, \sqrt{\delta}\right),$$

which, thanks to (2.4), in the normal form variables satisfy

$$\left(x_1^0, y_1^0, x_2^0, y_2^0\right) \sim \left(\delta\ln^{2^{j-1}}(1/\delta), 1, \sqrt{\delta}, \sqrt{\delta}\right).$$

Then, proceeding as before, these points are mapped to points of the form

$$\left(x_1^f, y_1^f, x_2^f, y_2^f\right) \sim \left(\sqrt{\delta}\ln^{2^{j-1}}(1/\delta), \delta^{1/2}, 1, \delta\ln(1/\delta)\right)$$

which in the original variables read

$$\mathcal{B}_1\left(p_1^0, q_1^0, p_2^0, q_2^0\right) \sim \left(\sqrt{\delta}\ln^{2^{j-1}}(1/\delta), \sqrt{\delta}, 1, \delta\ln^{2^j}(1/\delta)\right).$$

That is, the amount of logarithms doubles at each step and thus grows exponentially. This accumulation of logarithmic terms leads to *very bad estimates.* Indeed, to keep track of the orbit after $N-3$ local maps, we would need that

$$\delta\ln^{2^{N-3}}(1/\delta) \ll 1.$$

Therefore, we would need to choose δ extremely small with respect to N.

For example, if $\delta \gtrsim C^{-\mathcal{K}^{2a}} \sim C^{-2^{aN}}$ for some $C > 0$ independent of N, then the above expression gives

$$C^{-2^{aN}}\left(2^{aN}\ln C\right)^{2^{N-3}} \gg 1 \text{ for } a \leq 1.$$

In this case, the constant λ appearing in Theorem 4 would need to satisfy $\lambda \sim \delta^{-b}$ for some $b > 0$ and independent of N. As a result, Theorem 3 would give a diffusion time $T \sim \lambda^2\mathbb{K}\gamma N\ln 1/\delta \gtrsim C^{\mathcal{K}^2}$ (see formula (3.16)). Thus, choosing such a small δ would lead to very bad estimates for the diffusion time of Sobolev norms as we pointed out in Remark 1.4.

To overcome this problem, we modify slightly the initial conditions. Notice that if we choose x_1^0 such that

$$x_1^0 - 2y_1^0(x_2^0)^2\ln\sqrt{\delta} = 0,$$

we obtain that at the end $x_1^f \sim \sqrt{\delta}$ and thus we avoid the logarithmic term. This cancelation *will be crucial* in our proof. If we restrict x_1^0 to this set, we are taking $x_1^0 \sim \delta \ln(1/\delta)$ and therefore we will be sending points

$$\left(x_1^0, y_1^0, x_2^0, y_2^0\right) \sim \left(\delta \ln(1/\delta), 1, \sqrt{\delta}, \sqrt{\delta}\right),$$

to points

$$\left(x_1^f, y_1^f, x_2^f, y_2^f\right) \sim \left(\sqrt{\delta}, \sqrt{\delta}, 1, \delta \ln(1/\delta)\right).$$

The map will keep the same form expressed in the original variables, and, therefore, we will avoid having increasing separation from the invariant manifolds.

Note that for the true Toy Model is not integrable and therefore, we do not have a closed form for the flow near the saddle. Therefore, we need a very precise knowledge of the first orders of such dynamics so that we can impose analgous cancellations to the ones just explained to avoid deviation from the invariant manifolds. This knowledge is obtained by using the techniques developed by Shilnikov [Šil67] to analyze the local dynamics close to saddles which are resonant and therefore not well approximated by its linearization. Roughly speaking, for these systems, the linear part is not a good first order and if one considers the full nonlinear part the system is not integrable and therefore hard to analyze. Thus, one considers an intermediate first order in a clever way incorporating only some nonlinear terms. In this way, one obtains a good first order for this sytem simple enough to be analyzed. Therefore, one can obtain a precisely enough knowledge of the dynamics around the saddle to impose the explained cancellations. This is explained in more detail in Section 5, more precisely, in Lemma 5.2 and Remark 5.3.

2.3 Outline of the Proof

- Find symplectic coordinates near the origin in ℓ^1, where the original Hamiltonian $\mathcal{H}$ simplifies (see Theorem 2). Namely, $\mathcal{H} \circ \Gamma = \mathcal{D} + \widetilde{\mathcal{G}} + \mathcal{R}$, where $\mathcal{D}$ is a quadratic Hamiltonian, $\widetilde{\mathcal{G}}$ is of degree four and only contains resonant terms, and $\mathcal{R}$ is smaller.
- The dynamics of $\mathcal{D} + \widetilde{\mathcal{G}}$ has invariant finite-dimensional subspaces, which give rise to a simpler (no simple!) finite-dimensional Hamiltonian $h(b)$ given by (3.13). In terminology of [CKS+10] this Hamiltonian defines the *Toy Model.* In Theorem 3 we obtain orbits of the Toy Model which have transfer of energy.
- We show that are solutions of the system associated to $\mathcal{H}$ which are close to those of the Toy Model for long enough time (Theorem 4). These orbits undergo the wanted growth of the Sobolev norm.
- The proof of Theorem 3 occupies most of the paper. Theorems 2 and 4 are deferred to Appendices A and B respectively. Now we describe the plan of the proof of Theorem 3.
- Following [CKS+10] we detect a collection of periodic orbits $\{\mathbb{T}_j\}_{j=1}^{N-1}$ of $h(b)$, defined in (4.2), and heteroclinic orbits $\{\gamma_j\}_{j=1}^{N-2}$ connecting them (see (4.3)).

  The whole proof consists in a careful analysis of dynamics near the union of these periodic orbits and their connecting orbits. Our analysis naturally splits into

  - *local dynamics* near periodic orbits $\{\mathbb{T}_j\}_{j=1}^{N-1}$ and
  - *global dynamics* near heteroclinic orbits $\{\gamma_j\}_{j=1}^{N-2}$.

- More formally, Theorem 3 follows from Theorem 5. The latter Theorem in turn follows from Lemmas 4.7 and 4.8.
- The Local Lemma 4.7 provides refined information about the local behavior near the periodic orbits $\{\mathbb{T}_j\}_j$ with quantitative estimates.
- Global Lemma 4.8 provides refined information about the local behavior near the heteroclinic orbits from (4.3) with quantitative estimates.
- The proof of the Local Lemma 4.7 consists of several steps. As we have explained in Section 2.1, the periodic orbits $\{\mathbb{T}_j\}_j$ are of mixed type. Namely, in some directions the local behavior is hyperbolic, while in others it is elliptic. It turns out that the closer the orbits under investigation pass to the periodic orbits $\{\mathbb{T}_j\}_j$, the more decoupled (direct product-like) behavior they have.
- In Section 5 we set all the elliptic variables zero and study the (4-dimensional) Hyperbolic Toy Model.
- In Section 6 we use these results to deal with the full hyperbolic-elliptic system and prove Lemma 4.7.
- In Section 7 we prove the Global Lemma 4.8. As we pointed out, this implies Theorem 5, which in turn, implies Theorem 3.
- Combining this result with Theorem 2, proved in Appendix A, and Theorem 4, proved in Appendix B, we complete the proof of the main result (Theorem 1).

We summarize this in the following diagram:

$$\begin{array}{c}
\boxed{\text{Theorem 1}} \\
\Uparrow \\
\overbrace{\boxed{\text{Theorem 2}} + \underbrace{\boxed{\text{Theorem 3}}} + \boxed{\text{Theorem 4}}} \\
\Uparrow \\
\boxed{\text{Theorem 5}} \\
\Uparrow \\
\overbrace{\boxed{\text{Local Lemma 4.7}} + \boxed{\text{Global Lemma 4.8}}}
\end{array} \tag{2.5}$$

2.4 Major ingredients of the proof

We summarize here the new set of tools that we apply to the problem compared to [CKS$^+$10].

- In Theorem 2, we use a standard normal form (e.g. see [KP96]).
- Theorem 3 requires several new ideas:
  - Finitely smooth resonant normal form for hyperbolic saddles [BK94].
  - Shilnikov boundary value problem [Šil67] to study the local behavior close to the periodic orbits $\mathbb{T}_j$.

– As we explained for the model case in Section 2.2, to control the dynamics of the Toy Model we need a peculiar cancellation (see Section 5).
– To have cancellations at each stage, we need to establish local product structure for the orbits we are interested in (see Definition 4.3).

- Due to the good control of the solutions of the Toy Model, we are able to approximate the solutions of the original systems with the ones of the Toy Model for longer time compared with [CKS$^+$10] (see Theorem 4). To achieve this, we also modify the set Λ (see condition 6_Λ). This modification allows to slow down the spreading of mass outside Λ. This is explained in more detail in Appendix B

3 The three key theorems

We start the proof analyzing the infinite system of equations which describe the behavior of Fourier coefficients. Namely, consider the Fourier series of u,

$$u(t,x) = \sum_{n\in\mathbb{Z}^2} a_n(t)e^{inx}, \qquad a_n(t) := \hat{u}(t,n).$$

Therefore, equation (1.1) becomes an infinite system of equations for $\{a_n\}_{n\in\mathbb{Z}^2}$, which are given by

$$-i\dot{a}_n = |n|^2 a_n + \sum_{\substack{n_1,n_2,n_3\in\mathbb{Z}^2\\ n_1-n_2+n_3=n}} a_{n_1}\overline{a_{n_2}}a_{n_3}. \tag{3.1}$$

Note that this equation is Hamiltonian. Indeed, it can be written as

$$\dot{a}_n = 2i\partial_{\overline{a_n}}\mathcal{H}(a,\overline{a}),$$

where

$$\mathcal{H}(a,\overline{a}) = \mathcal{D}(a,\overline{a}) + \mathcal{G}(a,\overline{a}) \tag{3.2}$$

with

$$\mathcal{D}(a,\overline{a}) = \frac{1}{2}\sum_{n\in\mathbb{Z}^2}|n|^2|a_n|^2$$

$$\mathcal{G}(a,\overline{a}) = \frac{1}{4}\sum_{\substack{n_1,n_2,n_3,n_4\in\mathbb{Z}^2\\ n_1-n_2+n_3=n_4}} a_{n_1}\overline{a_{n_2}}a_{n_3}\overline{a_{n_4}}.$$

We will study equation (3.1) in a family of Banach spaces: all H^s-Sobolev spaces with $s>1$ as well as in the ℓ^1-space. The ℓ^1 space is defined as

$$\ell^1 = \left\{ a:\mathbb{Z}^2\to\mathbb{C} : \|a\|_{\ell^1} = \sum_{n\in\mathbb{Z}^2}|a_n| < \infty \right\}.$$

Note that, ℓ^1 is a Banach algebra with respect to the convolution product. Namely, if $a,b\in\ell^1$ its convolution product $a*b$, which is defined by

$$(a*b)_n = \sum_{n_1+n_2=n} a_{n_1}b_{n_2}$$

satisfies

$$\|a * b\|_{\ell^1} \leq \|a\|_{\ell^1}\|b\|_{\ell^1}. \tag{3.3}$$

Finally, let us point out that the L^2-norm conservation of (1.1), becomes now conservation of the ℓ^2-norm of a, defined as above. Namely, we have that $\|a(t)\|_{\ell^2} = \|a(0)\|_{\ell^2}$ for all $t \in \mathbb{R}$.

We want to study the evolution of certain solutions of equation (3.1), which will be small in the ℓ^1 norm. Now we make an outline of the proof.

The first step is to find out which terms make the biggest contribution to this evolution. To this end, we perform one step of normal form and bound the remainder in the ℓ^1-norm. We consider a small ball centered at the origin,

$$B(\eta) = \left\{a \in \ell^1 : \|a\|_{\ell^1} \leq \eta\right\}.$$

Theorem 2. *There exists $\eta > 0$ small enough such that there exists a symplectic change of coordinates $\Gamma : B(\eta) \to B(2\eta) \subset \ell^1$, $a = \Gamma(\alpha)$, which takes the Hamiltonian $\mathcal{H}$ in* (3.2) *into its Birkhoff normal form up to order four, that is,*

$$\mathcal{H} \circ \Gamma = \mathcal{D} + \widetilde{\mathcal{G}} + \mathcal{R},$$

where $\widetilde{\mathcal{G}}$ only contains resonant terms, namely

$$\widetilde{\mathcal{G}}(\alpha, \overline{\alpha}) = \frac{1}{4} \sum_{\substack{n_1,n_2,n_3,n_4 \in \mathbb{Z}^2 \\ n_1 - n_2 + n_3 = n_4 \\ |n_1|^2 - |n_2|^2 + |n_3|^2 = |n_4|^2}} \alpha_{n_1}\overline{\alpha_{n_2}}\alpha_{n_3}\overline{\alpha_{n_4}}$$

and $X_{\mathcal{R}}$, the vector field associated to the Hamiltonian $\mathcal{R}$, satisfies

$$\|\mathcal{X}_{\mathcal{R}}\|_{\ell^1} \leq \mathcal{O}\left(\|\alpha\|_{\ell^1}^5\right).$$

Moreover, the change Γ satisfies

$$\|\Gamma - \mathrm{Id}\|_{\ell^1} \leq \mathcal{O}\left(\|\alpha\|_{\ell^1}^3\right).$$

The proof of this theorem is postponed to Appendix A.

Once we perform one step of normal form, we have a new vector field

$$-i\dot{\alpha}_n = |n|^2\alpha_n + \sum_{(n_1,n_2,n_3) \in \mathcal{A}_0(n)} \alpha_{n_1}\overline{\alpha_{n_2}}\alpha_{n_3} + \partial_{\overline{\alpha_n}}\mathcal{R}, \tag{3.4}$$

where

$$\begin{aligned} \mathcal{A}_0(n) = \Big\{(n_1, n_2, n_3) \in \left(\mathbb{Z}^2\right)^3 :\ & n_1 - n_2 + n_3 = n, \\ & |n_1|^2 - |n_2|^2 + |n_3|^2 = |n|^2\Big\}. \end{aligned} \tag{3.5}$$

As a first step, we focus our attention to the degree 4 truncation of it, which will give the main contribution to the dynamics. Namely, we consider the Hamiltonian

$$\mathcal{H}' = \mathcal{D} + \widetilde{\mathcal{G}},$$

which has associated equations

$$-i\dot{\alpha}_n = |n|^2\alpha_n + \sum_{(n_1,n_2,n_3)\in\mathcal{A}_0(n)} \alpha_{n_1}\overline{\alpha_{n_2}}\alpha_{n_3}. \tag{3.6}$$

Note that the ℓ^2-norm of α is a first integral of this system as well as for (3.1) and (3.4). Namely,

$$\|\alpha(t)\|_{\ell^2} = \|\alpha(0)\|_{\ell^2} \text{ for all } t \in \mathbb{R}.$$

Then, to study the dynamics of α close to the origin (in the ℓ^1-norm) we remove its linear terms using the variation of constants formula. Moreover, we also remove certain cubic terms using the gauge freedom of equation (1.1). To this end, we make the change of coordinates

$$\alpha_n = \beta_n e^{i\left(G+|n|^2\right)t}, \tag{3.7}$$

where $G \in \mathbb{R}$ is a constant to be determined. The equations for β read

$$-i\dot{\beta}_n = -G\beta_n + \sum_{(n_1,n_2,n_3)\in\mathcal{A}_0(n)} \beta_{n_1}\overline{\beta_{n_2}}\beta_{n_3}.$$

Choosing G properly we can remove certain terms in the sum. Indeed, we split the sum as

$$\sum_{(n_1,n_2,n_3)\in\mathcal{A}_0(n)} = \sum_{\substack{(n_1,n_2,n_3)\in\mathcal{A}_0(n)\\ n_1,n_3\neq n}} + \sum_{\substack{(n_1,n_2,n_3)\in\mathcal{A}_0(n)\\ n_1=n}} + \sum_{\substack{(n_1,n_2,n_3)\in\mathcal{A}_0(n)\\ n_3=n}} - \sum_{\substack{(n_1,n_2,n_3)\in\mathcal{A}_0(n)\\ n_1=n_3=n}}$$

The last sum is just one term, which is given by $-\beta_n|\beta_n|^2$. The second and third sums, are in fact single sums and each of them is given by

$$\beta_n \sum_{k\in\mathbb{Z}^2} |\beta_k|^2 = \beta_n\|\beta\|_{\ell^2}^2.$$

Recall that both (3.6) and (3.7) preserve the ℓ^2-norm. Therefore, taking $G = 2\|\alpha\|_{\ell^2}^2 = 2\|\beta\|_{\ell^2}^2$, we can remove these two terms. Thus, with this choice, we obtain the equation for β, which reads

$$-i\dot{\beta}_n = -\beta_n|\beta_n|^2 + \sum_{n_1,n_2,n_3\in\mathcal{A}(n)} \beta_{n_1}\overline{\beta_{n_2}}\beta_{n_3} \tag{3.8}$$

where

$$\mathcal{A}(n) = \Big\{(n_1,n_2,n_3) \in \left(\mathbb{Z}^2\right)^3 : n_1 - n_2 + n_3 = n \\ |n_1|^2 - |n_2|^2 + |n_3|^2 = |n|^2, n_1 \neq n, n_3 \neq n\Big\}.$$

We define also the set of all resonant frequencies as

$$\mathcal{A} = \Big\{(n_1,n_2,n_3,n_4) \in \left(\mathbb{Z}^2\right)^4 : (n_1,n_2,n_3) \in \mathcal{A}(n_4)\Big\}.$$

Note that if $(n_1,n_2,n_3,n_4) \in \mathcal{A}$, then the four points form a rectangle in $\mathbb{Z}^2$ with the vertices ordered cyclically.

We reduce this system to a finite-dimensional one, which corresponds to an invariant finite-dimensional plane. To this end, we consider a set $\Lambda \subset \mathbb{Z}^2$ such that the corresponding harmonics do not interact with the harmonics outside of Λ. Moreover, we obtain a set Λ such that the harmonics in Λ interact in a very particular way. This set was constructed in [CKS$^+$10]. We explain now its construction and impose an additional condition on Λ from [CKS$^+$10].

Fix $N \gg 1$. Following [CKS$^+$10], we define a set $\Lambda \subset \mathbb{Z}^2$ consisting of N pairwise disjoint *generations*:

$$\Lambda = \Lambda_1 \cup \ldots \cup \Lambda_N.$$

Define a *nuclear family* to be a rectangle $(n_1, n_2, n_3, n_4) \in \mathcal{A}$ whose vertices are order, such that n_1 and n_3 (known as the *parents*) belong to a generation Λ_j and n_2 and n_4 (known as the *children*) live in the next generation Λ_{j+1}. Note that if (n_1, n_2, n_3, n_4) is a nuclear family, then so are (n_1, n_4, n_3, n_2), (n_3, n_2, n_1, n_4) and (n_3, n_4, n_1, n_2). These families are called trivial permutations of the family (n_1, n_2, n_3, n_4).

The conditions to impose to the set Λ are

1_Λ *Closure* If $n_1, n_2, n_3 \in \Lambda$ and $(n_1, n_2, n_3) \in \mathcal{A}(n)$, then $n \in \Lambda$. In other words, if three vertices of a rectangle are in Λ so is the fourth one.

2_Λ *Existence and uniqueness of spouse and children* For any $1 \leq j < N$ and any $n_1 \in \Lambda_j$, there exists a unique nuclear family (n_1, n_2, n_3, n_4) (up to trivial permutations) such that n_1 is a parent of this family. In particular, each $n_1 \in \Lambda_j$ has a unique spouse $n_3 \in \Lambda_j$ and has two unique children $n_2, n_4 \in \Lambda_{j+1}$ (up to permutation).

3_Λ *Existence and uniqueness of sibling and parents* For any $1 \leq j < N$ and any $n_2 \in \Lambda_{j+1}$, there exists a unique nuclear family (n_1, n_2, n_3, n_4) (up to trivial permutations) such that n_2 is a child of this family. In particular each $n_2 \in \Lambda_{j+1}$ has a unique sibling $n_4 \in \Lambda_{j+1}$ and two unique parents $n_1, n_3 \in \Lambda_j$ (up to permutation).

4_Λ *Nondegeneracy* The sibling of a frequency n is never equal to its spouse.

5_Λ *Faithfulness* Apart from the nuclear families, Λ does not contain any other rectangle.

These are the conditions imposed on Λ in [CKS$^+$10]. We will impose an additional condition:

6_Λ *No spreading condition* Let us consider any $n \notin \Lambda$. Then, n is vertex of at most two rectangles having two vertices in Λ and two vertices out of Λ.

Proposition 3.1. *Let $\mathcal{K} \gg 1$. Then, there exists $N \gg 1$ large and a set $\Lambda \subset \mathbb{Z}^2$, with*

$$\Lambda = \Lambda_1 \cup \ldots \cup \Lambda_N,$$

which satisfies conditions 1_Λ – 6_Λ and also

$$\frac{\sum_{n \in \Lambda_{N-1}} |n|^{2s}}{\sum_{n \in \Lambda_3} |n|^{2s}} \geq \frac{1}{2} 2^{(s-1)(N-4)} \geq \mathcal{K}^2. \tag{3.9}$$

Moreover, given any $R > 0$ (which may depend on $\mathcal{K}$), we can ensure that each generation Λ_j has 2^{N-1} disjoint frequencies n satisfying $|n| \geq R$.

The proof of Proposition 2.1 from [CKS$^+$10] applies except for proving that Condition 6_Λ is fullfilled, since this condition was not imposed in that paper. In Appendix C, we prove a quantitative version of this proposition and we show that modifying slightly the construction in [CKS$^+$10], one can construct a set Λ satisfying condition 6_Λ.

We use the set Λ to obtain a finite dimensional dynamical system (of high dimension) approximating (3.8). To this end, let us first note that, by Property 1_Λ, the manifold

$$M = \left\{\beta \in \mathbb{C}^{\mathbb{Z}^2} : \beta_n = 0 \text{ for all } n \notin \Lambda\right\}$$

is invariant by the flow associated to (3.8) and is finite dimensional. Indeed, by Proposition 3.1 its dimension is $N2^{N-1}$. Equation (3.8) restricted to M reads as follows. For each $n \in \Lambda$ we have

$$-i\dot{\beta}_n = -\beta_n|\beta_n|^2 + 2\beta_{n_{\mathrm{child}_1}}\beta_{n_{\mathrm{child}_2}}\overline{\beta_{n_{\mathrm{spouse}}}} + 2\beta_{n_{\mathrm{parent}_1}}\beta_{n_{\mathrm{parent}_2}}\overline{\beta_{n_{\mathrm{sibling}}}}. \tag{3.10}$$

Indeed, the presence of parents, children, and the sibling are guaranteed by 2_Λ and 3_Λ. Note, that in the first and last generations, the parents and children are set to zero respectively.

The manifold M has a submanifold of considerably lower dimension which is also invariant.

Corollary 3.2. *(cf. [CKS+10]) Consider the subspace*

$$\widetilde{M} = \{\beta \in M : \beta_{n_1} = \beta_{n_2} \text{ for all } n_1, n_2 \in \Lambda_j \text{ for some } j\},$$

where all the members of a generation take the same value. Then, $\widetilde{M}$ is invariant under the flow associated to (3.10)*.*

The dimension of $\widetilde{M}$ is equal to the number of generations, namely N. To define equation (3.10) restricted to $\widetilde{M}$, let us define

$$b_j = \beta_n \quad \text{for any } n \in \Lambda_j. \tag{3.11}$$

Then, (3.10) restricted to $\widetilde{M}$ becomes

$$\dot{b}_j = -ib_j^2\overline{b}_j + 2i\overline{b}_j\left(b_{j-1}^2 + b_{j+1}^2\right),\ j = 0, \dots N, \tag{3.12}$$

which is a Hamiltonian system with respect to the Hamiltonian

$$h(b) := \frac{1}{4}\sum_j |b_j|^4 - \frac{1}{2}\sum_j \left(\overline{b}_j^2 b_{j-1}^2 + b_j^2\overline{b}_{j-1}^2\right) \tag{3.13}$$

and the symplectic form $\Omega = \frac{i}{2} db_j \wedge d\overline{b}_j$.

Theorem 3. *Fix a large $\gamma \gg 1$. Then for any large enough N and $\delta = e^{-\gamma N}$, there exists an orbit of system* (3.12)*, $\nu > 0$ and $T_0 > 0$ such that*

$$\begin{array}{lll} |b_3(0)| > 1 - \delta^\nu & & |b_{N-1}(T_0)| > 1 - \delta^\nu \\ |b_j(0)| < \delta^\nu \quad \textit{for } j \neq 3 & \textit{and} & |b_j(T_0)| < \delta^\nu \quad \textit{for } j \neq N-1. \end{array}$$

Moreover, there exists a constant $\mathbb{K} > 0$ independent of N such that T_0 satisfies

$$0 < T_0 < \mathbb{K}N \ln\left(\frac{1}{\delta}\right) = \mathbb{K}\,\gamma\,N^2. \tag{3.14}$$

Remark 3.3. *An analog of this proposition also holds for some smaller δ, e.g. $\delta = C^{-2^N}$. This is related to Remark 1.4 about time of diffusion without cancelations.*

Using (3.11), Theorem 3 gives an orbit for equation (3.8). Moreover, both equations (3.8) and (3.12) are invariant under certain rescaling. Indeed if $b(t)$ is a solution of (3.12),

$$b^{\lambda}(t) = \lambda^{-1} b\left(\lambda^{-2} t\right) \tag{3.15}$$

is a solution of the same equation. By Theorem 3 duration of this solution in time is

$$T = \lambda^2 T_0 \leq \lambda^2 \, \mathbb{K} \, \gamma \, N^2, \tag{3.16}$$

where T_0 is the time obtained in Theorem 3, which satisfies (3.14).

We will see that, modulo a rotation of the modes (see (3.7)), there is a solution of equation (3.4) which is close to the orbit β^{λ} of (3.8) defined as

$$\begin{aligned} \beta_n^{\lambda}(t) &= \lambda^{-1} b_j\left(\lambda^{-2} t\right) \quad \text{for each} \quad n \in \Lambda_j \\ \beta_n^{\lambda}(t) &= 0 \quad \text{for each} \quad n \notin \Lambda. \end{aligned} \tag{3.17}$$

To have the original system being well approximated by the truncated system, we need that λ is large enough. Then the cubic terms in (3.4) dominate over the quintic ones. Nevertheless, the bigger λ, the slower the instability time by (3.16). Thus, we look for the smallest λ (with respect to N) for which the following approximation theorem applies.

Theorem 4. *Let* $\alpha(t) = \{\alpha_n(t)\}_{n \in \mathbb{Z}^2}$ *be the solution of* (3.4), $\beta^{\lambda}(t) = \{\beta_n^{\lambda}(t)\}_{n \in \mathbb{Z}^2}$ *be the solution of (3.8) given by* (3.17) *and* T *be the time defined in* (3.16). *Suppose* $\operatorname{supp} \alpha(0) \subset \Lambda$ *and* $\alpha(0) = \beta^{\lambda}(0)$. *Then, there exists a constant* $\kappa > 0$ *independent of* N *and* γ *such that, for*

$$\lambda = e^{\kappa \gamma N}, \tag{3.18}$$

and $0 < t < T$ *we have*

$$\sum_{n \in \mathbb{Z}^2} \left| \alpha_n(t) - e^{i(G + |n|^2)t} \beta_n^{\lambda}(t) \right| \leq \frac{1}{8} \lambda^{-2}, \tag{3.19}$$

where $G = 2\|\alpha(0)\|_{\ell^2}^2$.

Using the three key theorems: Theorems 2, 3 and 4 we complete the proof of Theorem 1.

Proof of Theorem 1. Using the change of variables Γ obtained in Theorem 2, from the solution α obtained in Theorem 4 we define $a = \Gamma(\alpha)$, which is a solution of system (3.1). We show that this orbit has the properties stated in Theorem 1.

To compute the growth of Sobolev norm of this orbit a, we use the notation

$$S_j = \sum_{n \in \Lambda_j} |n|^{2s} \text{ for } j = 1, \ldots, N-1. \tag{3.20}$$

To estimate the mass of our solution recall that $2^{N-1} = \sum_{n \in \Lambda_j} 1 = |\Lambda_j|$. We want to prove that

$$\frac{\|a(T)\|_{H^s}}{\|a(0)\|_{H^s}} \gtrsim \mathcal{K}$$

and estimate the mass $\|a(0)\|_{L^2}$ of the solution. To this end, we start by bounding $\|a(T)\|_{H^s}$ in terms of S_{N-1}. Since

$$\|a(T)\|_{H^s}^2 \geq \sum_{n \in \Lambda_{N-1}} |n|^{2s} \left|a_n(T)\right|^2 \geq S_{N-1} \inf_{n \in \Lambda_{N-1}} \left|a_n(T)\right|^2,$$

it is enough to obtain a lower bound for $|a_n(T)|$ with $n\in\Lambda_{N-1}$. Using the results of Theorems 2 and 4, we obtain

$$\begin{aligned}|a_n(T)| &\geq |\alpha_n(T)| - |\Gamma_n(\alpha)(T)-\alpha_n(T)|\\ &\geq \left|\beta_n^\lambda(T)e^{i\left(|n|^2+G\right)T}\right| - \left|\alpha_n(T)-\beta_n^\lambda(T)e^{i\left(|n|^2+G\right)T}\right|\\ &\quad - |\Gamma_n(\alpha)(T)-\alpha_n(T)|\,.\end{aligned} \tag{3.21}$$

We need to obtain a lower bound for the first term of the right hand side and upper bounds for the second and third ones. Indeed, using the definition of β^λ in (3.17) and the results in Theorem 3 we have that for $n\in\Lambda_{N-1}$,

$$\left|\beta_n^\lambda(T)\right|^2 = \lambda^{-2}\,|b_{N-1}(T_0)|^2 \geq \frac{3}{4}\lambda^{-2},$$

(the relation between T and T_0 is established in (3.16)).

For the second term in the right hand side of (3.21), it is enough to use Theorem 4 to obtain,

$$\left|\alpha_n(T)-\beta_n^\lambda(T)e^{i\left(|n|^2+G\right)T}\right|^2 \leq \left(\sum_{n\in\mathbb{Z}^2}\left|\alpha_n(T)-\beta_n^\lambda(T)e^{i\left(|n|^2+G\right)T}\right|\right)^2 \leq \frac{\lambda^{-2}}{8}.$$

For the lower bound of the third term, we use the bound for $\Gamma-\mathrm{Id}$ given in Theorem 2. Then,

$$|\Gamma_n(\alpha)(T)-\alpha_n(T)|^2 \leq \|\Gamma(\alpha)-\alpha\|_{\ell^1}^2 \leq \frac{\lambda^{-2}}{8}.$$

Thus, we can conclude that

$$\|\alpha(T)\|_{H^s}^2 \geq \frac{\lambda^{-2}}{2}S_{N-1}. \tag{3.22}$$

Now we prove that

$$\|a(0)\|_{H^s}^2 \lesssim \lambda^{-2}S_3 \qquad \text{and} \qquad \|a(0)\|_{L^2}^2 \lesssim \lambda^{-2}\,2^N. \tag{3.23}$$

By the definition of λ in (3.18), the second inequality implies that the mass of $a(0)$ is small. On the contrary, the first inequality does not imply that the H^s-norm of $a(0)$ is small. As a matter of fact S_3 is large[5].

To prove the first inequality of (3.23), let us point out that

$$\|a(0)\|_{H^s}^2 \leq \sum_{n\in\mathbb{Z}^2}|n|^{2s}\,|\alpha_n(0)+(\Gamma_n(\alpha(0)-\alpha_n(0))|^2\,.$$

We first bound $\|\alpha(0)\|_{H^s}^2$. To this end, let us recall that supp $\alpha=\Lambda$. Then, recalling also that $\alpha_n(0)=\beta_n^\lambda(0)$ (see Theorem 4), we have that

$$\|\alpha(0)\|_{H^s}^2 = \sum_{n\in\Lambda}|n|^{2s}\,|\alpha_n(0)|^2 = \sum_{n\in\Lambda}|n|^{2s}\left|\beta_n^\lambda(0)\right|^2.$$

[5]As pointed out to us by Terence Tao.

Recalling the definition of β^λ in (3.17) and the results in Theorem 3,

$$\begin{aligned}\sum_{n\in\Lambda}|n|^{2s}\left|\beta_n^\lambda(0)\right|^2 &\leq (1-\delta^\nu)\,S_3+\delta^\nu\sum_{j\neq 3}S_j\\ &\leq S_3\left(1-\delta^\nu+\delta^\nu\sum_{j\neq 3}\frac{S_j}{S_3}\right).\end{aligned}$$

From Proposition 3.1 we know that $j\neq 3$,

$$\frac{S_j}{S_3}\lesssim e^{sN}$$

Therefore, to bound these terms we use the definition of δ from Theorem 3 taking $\gamma=\widetilde{\gamma}(s-1)$. Since $s-1>s_0-1>0$ is fixed, we can choose such $\widetilde{\gamma}\gg 1$. Then, we have that

$$\|\alpha(0)\|_{H^s}^2=\sum_{n\in\Lambda}|n|^{2s}\left|\beta_n^\lambda(0)\right|^2\lesssim\lambda^{-2}S_3.$$

To complete the proof of statement (3.23) recall that the support of $\Gamma(\alpha)-\alpha$ is

$$\Lambda^3=\left\{n\in\mathbb{Z}^2:n=n_1-n_2+n_3,n_1,n_2,n_3\in\Lambda\right\}$$

and apply Theorem 2.

Using inequalities (3.22) and (3.23), we have that

$$\frac{\|a(T)\|_{H^s}^2}{\|a(0)\|_{H^s}^2}\gtrsim\frac{S_{N-1}}{S_3},$$

and then, applying Proposition 3.1, we obtain

$$\frac{\|a(T)\|_{H^s}^2}{\|a(0)\|_{H^s}^2}\gtrsim\frac{1}{2}2^{(s-1)(N-4)}\geq\mathcal{K}^2.$$

It is left to estimate the diffusion time T. Use Proposition 3.1 to set $\mathcal{K}\simeq 2^{(s-1)N/2}$ and $c=4\kappa\gamma/(s-1)$, and definition (3.18) to set $\lambda=e^{\kappa\gamma N}\simeq\mathcal{K}^{c/(2\ln 2)}$. Then, for the time of diffusion we obtain

$$|T|\leq\mathbb{K}\;\gamma\;\lambda^2N^2\leq\mathbb{K}\;\gamma\;\mathcal{K}^{c/\ln 2}\frac{4\ln^2\mathcal{K}}{\ln^2 2\;(s-1)^2}\leq\mathcal{K}^c$$

for large $\mathcal{K}$. This completes the proof of Theorem 1. □

4 The finite dimensional model: proof of Theorem 3

We devote this section to describe the proof of Theorem 3. The proofs of the partial results stated in this section are deferred to Sections 5–7.

To prove Theorem 3 we need to analyze certain orbits of system (3.12) given by Hamiltonian $h(b)$ in (3.13). This system has another conserved quantity: the mass

$$\mathcal{M}(b)=\sum|b_j|^2. \tag{4.1}$$

We obtain the orbits given in Theorem 3 on the manifold $\mathcal{M}(b) = 1$.

It can be easily seen that on $\mathcal{M}(b) = 1$ there are periodic orbits $\mathbb{T}_j$ given by

$$b_j(t) = e^{-it}, \ b_k(t) = 0 \text{ for } k \neq j, \tag{4.2}$$

which in the normal directions are of mixed type: hyperbolic in some directions and elliptic in the others. Moreover, there exist two families of heteroclinic orbits, which connect consecutive periodic orbits. Consider the 2-dimensional complex plane $L_j = \{\forall k \neq j, j+1 : \ b_k = 0\}$. In Section 2.1 we show that they are invariant and the dynamics inside is integrable. Then, the (two dimensional) unstable manifold of the periodic orbit $(b_j(t), b_{j+1}(t)) = (e^{-it}, 0)$ coincides with the (two dimensional) stable manifold of $(b_j(t), b_{j+1}(t)) = (0, e^{-it})$ and it is foliated by heteroclinic orbits. As usual, the stable and unstable invariant manifolds have two branches and, therefore, we have two families of heteroclinic connections. It turns out that they can be explicitly computed [CKS$^+$10] and are given by

$$\gamma_j^{\pm}(t) = (0, \dots, 0, b_j(t), b_{j+1}^{\pm}(t), 0, \dots, 0) \tag{4.3}$$

with

$$b_j(t) = \frac{e^{-i(t+\vartheta)}\omega}{\sqrt{1 + e^{2\sqrt{3}t}}}, \qquad b_{j+1}^{\pm}(t) = \pm\frac{e^{-i(t+\vartheta)}\omega^2}{\sqrt{1 + e^{-2\sqrt{3}t}}}, \qquad \vartheta \in \mathbb{T}.$$

To prove Theorem 3 we look for an orbit which shadows the sequence of separatrices, as follows

- it starts close to the periodic orbit $\mathbb{T}_3$
- later it passes close to the periodic orbit $\mathbb{T}_4$
- later it passes close to the periodic orbit $\mathbb{T}_5$ and so on
- finally it arrives to a neighborhood of the periodic orbit $\mathbb{T}_{N-1}$.

Our main goal is to prove

existence of such orbits and estimate the transition time in terms of N.

In making these transition we have the freedom of whether to travel close to γ_j^+ or γ_j^-. We will choose always γ_j^+ The procedure for γ_j^- is analogous.

We believe it is helpful to the reader to have the following information about the transition of energy. We have a solution $b(t) = \{b_j(t)\}_{j=0,\dots,N}$ of the system (3.12). We fix $\sigma > 0$ small, but independent of N, and $\delta = e^{-\gamma N}$. For each $j = 2, \dots, N-1$ near the periodic orbit $\mathbb{T}_j$ and later near $\mathbb{T}_{j+1}$ we have the following table of orders of magnitude of distribution of energy

$$\begin{array}{ccc}
\text{near } \mathbb{T}_j & \longrightarrow & \text{near } \mathbb{T}_{j+1} \\
|b_{<j-2}| & \longrightarrow & |b_{<j-2}|\,(1 + O(\delta^{r'})) \\
|b_{j-2}| & \longrightarrow & K|b_{j-2}| \\
|b_{j-1}| = O(\sigma) & \longrightarrow & (C^{(j)}\delta)^{1/2} \\
|b_j| = 1 - O(\sigma^2) \text{ (mass conservation)} & \longrightarrow & O(\sigma) \\
|b_{j+1}| = (C^{(j)}\delta)^{1/2} & \longrightarrow & 1 - O(\sigma^2) \text{ (mass conservation)} \\
|b_{j+2}| & \longrightarrow & K|b_{j+2}| \\
|b_{>j+2}| & \longrightarrow & |b_{>j+2}|\,(1 + O(\delta^{r'})).
\end{array} \tag{4.4}$$

We decompose a diffusing orbit into $N-5$ parts: near each periodic orbit $\mathbb{T}_j,\ j=3,\dots,N-1$ we construct sections transversal to the flow so that they divide the orbit appropriately. For each transition from one section to the next one we associate a map $\mathcal{B}^j$ which sends points close to $\mathbb{T}_j$ to points close to $\mathbb{T}_{j+1}$. This leads to analysis of the composition of all these maps

$$\mathcal{B}^* = \mathcal{B}^{N-1} \circ \dots \circ \mathcal{B}^3.$$

To study these maps we will consider different systems of coordinates which, on one hand, will take advantage of the fact that mass (4.1) is a conserved quantity, and on the other hand, will be adapted to the linear normal behavior of the periodic orbits. These systems of coordinates are specified in Section 4.1.

4.1 Symplectic reduction and diagonalization

To study the different transition maps we use a system of coordinates defined in [CKS+10]. It consists of two steps:

- A symplectic reduction, which uses that mass (4.1) is conserved and sends the periodic orbit $\mathbb{T}_j$ into a critical point.
- A linear transformation which diagonalizes the linearization of dynamics near this critical point.

We perform the change corresponding to the traveling close to the j periodic orbit $\mathbb{T}_j$. We restrict ourselves to $\mathcal{M}(b)=1$ and we take

$$b_j = r^{(j)} e^{i\theta^{(j)}},\ b_k = c_k^{(j)} e^{i\theta^{(j)}} \quad \text{for all} \quad k \neq j, \tag{4.5}$$

where $\theta^{(j)}$ is a variable on $\mathbb{T}_j$. From now on in this section we omit the superscripts (j). It can be seen that after eliminating r using that $\mathcal{M}(b)=1$ and omitting the equation for the variable θ, one obtain a new set of equations whose c_k components form a Hamiltonian system with the Hamiltonian

$$\begin{aligned} H^{(j)}(c) =& \frac{1}{4}\sum_{k\neq j} |c_k|^4 + \frac{1}{4}\left(1-\sum_{k\neq j}|c_k|^2\right)^2 - \frac{1}{2}\sum_{k\neq j,j+1} \overline{c}_k^2 c_{k-1}^2 + c_k^2\overline{c}_{k-1}^2 \\ &- \frac{1}{2}\left(1-\sum_{k\neq j}|c_k|^2\right)\left(c_{j-1}^2+\overline{c}_{j-1}^2+c_{j+1}^2+\overline{c}_{j+1}^2\right) \end{aligned}$$

and the symplectic form $\Omega = \frac{i}{2} dc_k \wedge d\overline{c_k}$. The Hamiltonian $H^{(j)}(c)$ can be written as

$$H^{(j)}(c) = H_2^{(j)}(c) + H_4^{(j)}(c) \tag{4.6}$$

with

$$\begin{aligned} H_2^{(j)}(c) =& -\frac{1}{2}\sum_{k\neq j}|c_k|^2 - \frac{1}{2}\left(c_{j-1}^2+\overline{c}_{j-1}^2+c_{j+1}^2+\overline{c}_{j+1}^2\right) \\ H_4^{(j)}(c) =& \frac{1}{4}\sum_{k\neq j}|c_k|^4 + \frac{1}{4}\left(\sum_{k\neq j}|c_k|^2\right)^2 - \frac{1}{2}\sum_{k\neq j,j+1}\overline{c}_k^2 c_{k-1}^2 + c_k^2\overline{c}_{k-1}^2 \\ &+ \frac{1}{2}\sum_{k\neq j}|c_k|^2\left(c_{j-1}^2+\overline{c}_{j-1}^2+c_{j+1}^2+\overline{c}_{j+1}^2\right). \end{aligned} \tag{4.7}$$

Since we are omitting the evolution of the variable θ, the periodic orbit $\mathbb{T}_j$ has become now a critical point for the equation associated to this Hamiltonian, which is defined as $c = 0$. For the same reason, the two families of heteroclinic connections defined in (4.3), now have become just two one dimensional heteroclinic connections.

The second step is to look for a change of variables which diagonalizes the vector field around this critical point. This change only modifies the coordinates (c_{j-1}, c_{j+1}) and is given by

$$\begin{pmatrix} c_{j-1} \\ c_{j+1} \end{pmatrix} = \begin{pmatrix} \omega^2 p_1 + \omega q_1 \\ \omega^2 p_2 + \omega q_2 \end{pmatrix} \tag{4.8}$$

where $\omega = e^{2\pi i/3}$ (see [CKS+10]). Note that this change is conformal and leads to the symplectic form

$$\widetilde{\Omega} = \frac{i}{2} dc_k \wedge d\bar{c}_k + dp_1 \wedge dq_1 + dp_2 \wedge dq_2. \tag{4.9}$$

To study the Hamiltonian expressed in the new variables let us introduce some notation. We define

$$\mathcal{P}_j = \{1 \leq k \leq N; k \neq j-1, j, j+1\}, \tag{4.10}$$

which is the set of subindexes of the elliptic modes. From now on we will denote by q and p all the stable and unstable coordinates $q = (q_1, q_2)$ and $p = (p_1, p_2)$ respectively and by c all the elliptic modes, namely c_k with $k \in \mathcal{P}_j$.

Lemma 4.1. *The change* (4.8) *transforms the Hamiltonian* (4.6) *into the Hamiltonian*

$$\widetilde{H}^{(j)}(p,q,c) = \widetilde{H}_2^{(j)}(p,q,c) + \widetilde{H}_4^{j}(p,q,c) \tag{4.11}$$

with homogeneous polynomials

$$\widetilde{H}_2^{(j)}(p,q,c) = -\frac{1}{2}\sum_{k\in\mathcal{P}_j} |c_k|^2 + \sqrt{3}\,(p_1 q_1 + p_2 q_2)$$

and

$$\widetilde{H}_4^{(j)}(p,q,c) = \widetilde{H}_{\text{hyp}}^{(j)}(p,q) + \widetilde{H}_{\text{ell}}^{(j)}(c) + \widetilde{H}_{\text{mix}}^{(j)}(p,q,c)$$

where

$$\widetilde{H}_{\text{hyp}}^{(j)}(p,q) = \sum_{k=1}^{3} \nu_k p_1^k q_1^{4-k} + \sum_{k=1}^{3} \nu_k p_2^k q_2^{4-k} + \sum_{k,\ell=0}^{2} \nu_{k\ell} p_1^k q_1^{2-k} p_2^{\ell} q_2^{2-\ell}$$

$$\begin{aligned} \widetilde{H}_{\text{ell}}^{(j)}(c) =& \frac{1}{4}\sum_{k\in\mathcal{P}_j} |c_k|^4 + \frac{1}{4}\left(\sum_{k\in\mathcal{P}_j} |c_k|^2\right)^2 \\ &- \frac{1}{2}\sum_{k\neq j-1,j,j+1,j+2} c_k^2 \overline{c_{k-1}}^2 + \overline{c_k}^2 c_{k-1}^2 \end{aligned} \tag{4.12}$$

$$\begin{aligned} \widetilde{H}_{\text{mix}}^{(j)}(p,q,c) =& -\frac{\sqrt{3}}{2}\sum_{k\in\mathcal{P}_j} |c_k|^2 (q_1 p_1 + q_2 p_2) \\ &- \frac{1}{2}\left(\omega^2 p_1 + \omega q_1\right)^2 \overline{c_{j-2}}^2 - \frac{1}{2}\left(\omega^2 q_1 + \omega p_1\right)^2 c_{j-2}^2 \\ &- \frac{1}{2}\left(\omega^2 p_2 + \omega q_2\right)^2 \overline{c_{j+2}}^2 - \frac{1}{2}\left(\omega^2 q_2 + \omega p_2\right)^2 c_{j+2}^2 \end{aligned} \tag{4.13}$$

for certain constants and $\nu_k, \nu_{k\ell} \in \mathbb{R}$.

Remark 4.2. *Even though the proof of this lemma is a simple substitution of (p,q) we do need specifics of the form of the decomposition into Hamiltonians:*

- $\widetilde{H}_2^{(j)}$ *is the direct product of two linear saddles (p_i,q_i), $i=1,2$ and $N-2$ linear elliptic points $\{c_k\}_k$, $k\in\mathcal{P}_j$.*

- $\widetilde{H}_{\rm hyp}^{(j)}$ *consists only of some saddle terms. In particular, it does not contain terms p_i^4,q_i^4, $i=1,2$ so $\{q=0\}$ and $\{p=0\}$ are invariant manifolds of $\widetilde{H}$ if we set $c=0$. This implies that the two heteroclinic orbits which connect the critical point $(p,q,c)=(0,0,0)$ to the next periodic orbit $\mathbb{T}_{j+1}$ are just defined as*

$$\left(p_1^\pm(t),q_1^\pm(t),p_2^\pm(t),q_2^\pm(t),c^\pm(t)\right)=\left(0,0,\frac{\pm1}{1+e^{-2\sqrt{3}t}},0,0\right).$$

  Moreover, $\mathbb{T}_{j+1}$ is now defined as $|c_{j+1}|=1$. Due to (4.8) it is equivalent to $p_2^2+q_2^2-p_2q_2=1$.

- *Near $p=q=0$, which corresponds to the periodic orbit $\mathbb{T}_j$ Hamiltonians $\widetilde{H}_{\rm ell}^{(j)}$ and $\widetilde{H}_{\rm mix}^{(j)}$ are almost integrable. The only source of non-integrability comes from the second line of (4.12) for $\widetilde{H}_{\rm ell}^{(j)}$ and from the second and third line of (4.13) for $\widetilde{H}_{\rm mix}^{(j)}$.*

- *Later we select regions with c's being exponentially small in N. As a result, coupling between hyperbolic variables – (p,q) and elliptic ones c's is exponentially small in N. This decoupling at the leading order is crucial for our analysis.*

- *Among all the constants ν_k which appear in the definition of Hamiltonian (4.11), $\nu_{02}\neq0$ is the only one which plays a significant role in the proof of Theorem 3. Indeed, the corresponding term is resonant and will be the leading term in studying the transition close to the saddle. We assume, without loss of generality that $\nu_{02}>0$ since the case $\nu_{02}<0$ can be done analogously.*

Proof. To obtain the explicit form of $\widetilde{H}_4^{(j)}$, note that $H_4^{(j)}(c)$ in (4.7) can be rewritten as

$$\begin{aligned}H_4^{(j)}(c)=&\frac14\sum_{k\neq j}|c_k|^4+\frac14\left(\sum_{k\neq j}|c_k|^2+c_{j-1}^2+\bar{c}_{j-1}^2+c_{j+1}^2+\bar{c}_{j+1}^2\right)^2\\&-\frac12\sum_{k\neq j,j+1}\bar{c}_k^2c_{k-1}^2+c_k^2\bar{c}_{k-1}^2-\frac14\left(c_{j-1}^2+\bar{c}_{j-1}^2+c_{j+1}^2+\bar{c}_{j+1}^2\right)^2.\end{aligned}$$

Written in this way, the second term in the first row is just a constant times $\widetilde{H}_2^{(j)}$ squared. Then, the particular form of $\widetilde{H}_{\rm hyp}^{(j)}$, $\widetilde{H}_{\rm ell}^{(j)}$, and $\widetilde{H}_{\rm mix}^{(j)}$ can be obtained just performing the change of coordinates. □

Since the symplectic form is given by (4.9), equations associated to the Hamiltonian (4.11) are

$$\begin{aligned}
\dot p_1 &= \sqrt{3}p_1 + \mathcal{Z}_{\mathrm{hyp},p_1} + \mathcal{Z}_{\mathrm{mix},p_1} = \sqrt{3}p_1 + \partial_{q_1}\widetilde{H}^{(j)}_{\mathrm{hyp}} + \partial_{q_1}\widetilde{H}^{(j)}_{\mathrm{mix}}\\
\dot q_1 &= -\sqrt{3}q_1 + \mathcal{Z}_{\mathrm{hyp},q_1} + \mathcal{Z}_{\mathrm{mix},q_1} = -\sqrt{3}q_1 - \partial_{p_1}\widetilde{H}^{(j)}_{\mathrm{hyp}} - \partial_{p_1}\widetilde{H}^{(j)}_{\mathrm{mix}}\\
\dot p_2 &= \sqrt{3}p_2 + \mathcal{Z}_{\mathrm{hyp},p_2} + \mathcal{Z}_{\mathrm{mix},p_2} = \sqrt{3}p_2 + \partial_{q_2}\widetilde{H}^{(j)}_{\mathrm{hyp}} + \partial_{q_2}\widetilde{H}^{(j)}_{\mathrm{mix}}\\
\dot q_2 &= -\sqrt{3}q_2 + \mathcal{Z}_{\mathrm{hyp},q_2} + \mathcal{Z}_{\mathrm{mix},q_2} = -\sqrt{3}q_2 - \partial_{p_2}\widetilde{H}^{(j)}_{\mathrm{hyp}} - \partial_{p_2}\widetilde{H}^{(j)}_{\mathrm{mix}}\\
\dot c_k &= ic_k + \mathcal{Z}_{\mathrm{ell},c_k} + \mathcal{Z}_{\mathrm{mix},c_k} = ic_k - 2i\partial_{c_k}\widetilde{H}^{(j)}_{\mathrm{ell}} - 2i\partial_{c_k}\widetilde{H}^{(j)}_{\mathrm{mix}}.
\end{aligned} \tag{4.14}$$

where

$$\begin{aligned}
\mathcal{Z}_{\mathrm{hyp},p_1} =& \sum_{k=1}^{3}(4-k)\nu_k p_1^k q_1^{3-k} + \nu_{12}p_1p_2^2 + \nu_{11}p_1p_2q_2 + \nu_{10}p_1q_2^2\\
&+ 2\nu_{02}q_1p_2^2 + 2\nu_{01}q_1p_2q_2 + 2\nu_{00}q_1q_2^2
\end{aligned} \tag{4.15}$$

$$\begin{aligned}
\mathcal{Z}_{\mathrm{hyp},q_1} =& -\sum_{k=1}^{3}k\nu_k p_1^{k-1}q_1^{4-k} - 2\nu_{22}p_1p_2^2 - 2\nu_{21}p_1p_2q_2 - 2\nu_{20}p_1q_2^2\\
&- \nu_{12}q_1p_2^2 - \nu_{11}q_1p_2q_2 - \nu_{10}q_1q_2^2
\end{aligned} \tag{4.16}$$

$$\begin{aligned}
\mathcal{Z}_{\mathrm{hyp},p_2} =& \sum_{k=1}^{4}(4-k)\nu_k p_2^k q_2^{3-k} + \nu_{21}p_1^2p_2 + \nu_{11}p_1q_1p_2 + \nu_{01}q_1^2p_2\\
&+ 2\nu_{20}p_1^2q_2 + 2\nu_{10}p_1q_1q_2 + 2\nu_{00}q_1^2q_2
\end{aligned} \tag{4.17}$$

$$\begin{aligned}
\mathcal{Z}_{\mathrm{hyp},q_2} =& -\sum_{k=1}^{4}k\nu_{k\ell}p_2^{k-1}q_2^{4-k} - 2\nu_{22}p_1^2p_2 - 2\nu_{12}p_1q_1p_2 - 2\nu_{02}q_1^2p_2\\
&- \nu_{21}p_1^2q_2 - \nu_{11}p_1q_1q_2 - \nu_{01}q_1^2q_2
\end{aligned} \tag{4.18}$$

$$\mathcal{Z}_{\mathrm{ell},c_k} = -i|c_k|^2c_k - i\left(\sum_{\ell\in\mathcal{P}_j}|c_\ell|^2\right)c_k + 2i\overline{c_k}\left(c_{k-1}^2 + c_{k+1}^2\right) \tag{4.19}$$

$$\mathcal{Z}_{\mathrm{mix},q_1} = \omega^2(\omega^2p_1+\omega q_1)\overline{c_{j-2}}^2 + \omega(\omega p_1+\omega^2q_1){c_{j-2}}^2 + \frac{\sqrt{3}}{2}\sum_{\ell\in\mathcal{P}_j}|c_\ell|^2q_1 \tag{4.20}$$

$$\mathcal{Z}_{\mathrm{mix},p_1} = -\omega(\omega^2p_1+\omega q_1)\overline{c_{j-2}}^2 - \omega^2(\omega p_1+\omega^2q_1){c_{j-2}}^2 - \frac{\sqrt{3}}{2}\sum_{\ell\in\mathcal{P}_j}|c_\ell|^2p_1 \tag{4.21}$$

$$\mathcal{Z}_{\mathrm{mix},q_2} = \omega^2(\omega^2p_2+\omega q_2)\overline{c_{j+2}}^2 + \omega(\omega p_2+\omega^2q_2){c_{j+2}}^2 + \frac{\sqrt{3}}{2}\sum_{\ell\in\mathcal{P}_j}|c_\ell|^2q_2 \tag{4.22}$$

$$\mathcal{Z}_{\mathrm{mix},p_2} = -\omega(\omega^2p_2+\omega q_2)\overline{c_{j+2}}^2 - \omega^2(\omega p_2+\omega^2q_2){c_{j+2}}^2 - \frac{\sqrt{3}}{2}\sum_{\ell\in\mathcal{P}_j}|c_\ell|^2p_2 \tag{4.23}$$

$$\mathcal{Z}_{\mathrm{mix},c_k} = i\sqrt{3}c_k(q_1p_1+q_2p_2) \quad \text{for } k\in\mathcal{P}_j\setminus\{j\pm2\} \tag{4.24}$$

$$\begin{aligned}
\mathcal{Z}_{\mathrm{mix},c_{j-2}} =& i\sqrt{3}c_{j-2}(q_1p_1+q_2p_2) - 2i(\omega^2p_1+\omega q_1)^2\overline{c_{j-2}}\\
\mathcal{Z}_{\mathrm{mix},c_{j+2}} =& i\sqrt{3}c_{j+2}(q_1p_1+q_2p_2) - 2i(\omega^2p_2+\omega q_2)^2\overline{c_{j+2}}.
\end{aligned} \tag{4.25}$$

4.2 The iterative Theorem

Now that we have obtained the adapted coordinates for each saddle we are ready to explain the strategy to prove Theorem 3. To obtain the orbit given in Theorem 3, we will consider several co-dimension one sections $\{\Sigma_j^{\mathrm{in}}\}_{j=1}^N$ and transition maps $\mathcal{B}^j$ from one section Σ_j^{in} to the next one $\Sigma_{j+1}^{\mathrm{in}}$. Then, we will detect a class of open sets $\{\mathcal{V}_j\}_j$, $\mathcal{V}_j \subset \Sigma_j^{\mathrm{in}}$, $j = 1, \dots, N-1$, which have a certain *almost product structure* (see Definition 4.3) such that $\mathcal{V}_{j+1} \subset \mathcal{B}^j(\mathcal{V}_j)$ and none of them is empty. Each set $\mathcal{V}_j$ is located close to the stable manifold of the periodic orbit $\mathbb{T}_j$. Composing all these maps we will be able to find orbits claimed to exist in Theorem 3.

We start by defining these maps. The first step is to define certain transversal sections to the flow. We use the coordinates adapted to the saddle j, $(p^{(j)}, q^{(j)}, c^{(j)})$, which have been introduced in Section 4.1, to define these sections. Indeed, in these coordinates, it can be easily seen that the heteroclinic connections (4.3), which connect $(p^{(j)}, q^{(j)}, c^{(j)}) = (0,0,0)$ with the previous and next saddles are defined by $(q_1^{(j)}, p_2^{(j)}, q_2^{(j)}, c^{(j)}) = (0,0,0,0)$ and $(p_1^{(j)}, q_1^{(j)}, p_2^{(j)}, c^{(j)}) = (0,0,0,0)$ respectively. Thus, we define the map $\mathcal{B}^j$ from the section

$$\Sigma_j^{\mathrm{in}} = \left\{ q_1^{(j)} = \sigma \right\} \tag{4.26}$$

to the section

$$\Sigma_{j+1}^{\mathrm{in}} = \left\{ q_1^{(j+1)} = \sigma \right\}.$$

Here $\sigma > 0$ is a small parameter that will be determined later on. In fact, we do not define the map $\mathcal{B}^j$ in the whole section but in an open set $\mathcal{V}^j \subset \Sigma_j^{\mathrm{in}}$, which lies close to the heteroclinic that connects the saddle $j-1$ to the saddle j. Then, we will consider maps

$$\mathcal{B}^j : \mathcal{V}_j \subset \Sigma_j^{\mathrm{in}} \to \Sigma_{j+1}^{\mathrm{in}}$$

and we will choose the sets $\mathcal{V}^j$ recursively in such a way that

$$\mathcal{V}_{j+1} \subset \mathcal{B}^j(\mathcal{V}_j). \tag{4.27}$$

This condition will allow us to compose all the maps $\mathcal{B}^j$. Indeed, the domain of definition of the map $\mathcal{B}^{j+1}$ will intersect the image of the map $\mathcal{B}^j$ in an open set.

The sets $\mathcal{V}_j$ will have a product-like structure as is stated in the next definition. Before stating it, we introduce some notation. We define the subsets of indices $\mathcal{P}_j$ in (4.10),

$$\begin{aligned} \mathcal{P}_j^- &= \{k = 1, \dots, j-3\} \\ \mathcal{P}_j^+ &= \{k = j+3, \dots, N\}. \end{aligned} \tag{4.28}$$

The first set consists of preceding non-neighbor modes to $j-1$, the second — of foreseeing non-neighbor modes to $j+1$. The modes $k = j \pm 2$ are called *adjacent*. These modes have a stronger interaction with the hyperbolic modes.

Note that we split the non-neighbor elliptic modes in two sets: the $+$ stands for *future* $-$ stands for *past*. Indeed, along orbits we study future modes will eventually become hyperbolic in the future, past have already been hyperbolic. Analogously, we call future adjacent — the mode $c_{j+2}^{(j)}$ and past adjacent — $c_{j-2}^{(j)}$.

For a point $(p^{(j)}, q^{(j)}, c^{(j)}) \in \Sigma_j^{\mathrm{in}}$, we define $c_-^{(j)} = (c_1^{(j)}, \dots, c_{j-2}^{(j)})$ and $c_+^{(j)} = (c_{j+2}^{(j)}, \dots, c_N^{(j)})$. We define also the projections $\pi_\pm(p^{(j)}, q^{(j)}, c^{(j)}) = c_\pm^{(j)}$ and $\pi_{\mathrm{hyp},+} = (p^{(j)}, q^{(j)}, c_+^{(j)})$.

Definition 4.3. *Fix positive constants $r \in (0,1)$, δ and σ and define a multi-parameter set of positive constants*

$$\mathcal{I}_j = \left\{C^{(j)}, m_{\rm ell}^{(j)}, M_{{\rm ell},\pm}^{(j)}, m_{\rm adj}^{(j)}, M_{{\rm adj},\pm}^{(j)}, m_{\rm hyp}^{(j)}, M_{\rm hyp}^{(j)}\right\}. \tag{4.29}$$

Then, we say that a (non-empty) set $\mathcal{U} \subset \Sigma_j^{\rm in}$ has an $\mathcal{I}_j$-product-like structure if it satisfies the following two conditions:

C1

$$\mathcal{U} \subset \mathbb{D}_j^1 \times \ldots \times \mathbb{D}_j^{j-2} \times \mathcal{N}_j^+ \times \mathbb{D}_j^{j+2} \times \ldots \times \mathbb{D}_j^N,$$

where

$$\mathbb{D}_j^k = \left\{\left|c_k^{(j)}\right| \leq M_{{\rm ell},\pm}^{(j)} \delta^{(1-r)/2}\right\} \quad \text{for } k \in \mathcal{P}_j^\pm$$
$$\mathbb{D}_j^{j\pm2} \subset \left\{\left|c_{j\pm2}^{(j)}\right| \leq M_{{\rm adj},\pm}^{(j)} \left(C^{(j)}\delta\right)^{1/2}\right\}$$

and

$$\begin{aligned}\mathcal{N}_j^+ = \Big\{ &\left(p_1^{(j)}, q_1^{(j)}, p_2^{(j)}, q_2^{(j)}\right) \in \mathbb{R}^4 : \\ &-C^{(j)}\delta\left(\ln(1/\delta) + M_{\rm hyp}^{(j)}\right) \leq p_1^{(j)} \leq -C^{(j)}\delta\left(\ln(1/\delta) - M_{\rm hyp}^{(j)}\right), \\ &q_1^{(j)} = \sigma,\ g_{\mathcal{I}_j}(p_2, q_2, \sigma, \delta) = 0,\ |p_2^{(j)}|, |q_2^{(j)}| \leq M_{\rm hyp}^{(j)}\left(C^{(j)}\delta\right)^{1/2}\Big\}.\end{aligned} \tag{4.30}$$

C2

$$\mathcal{N}_j^- \times \mathbb{D}_{j,-}^{j+2} \times \ldots \times \mathbb{D}_{j,-}^N \subset \pi_{{\rm hyp},+}\mathcal{U},$$

where

$$\mathbb{D}_{j,-}^k = \left\{\left|c_k^{(j)}\right| \leq m_{\rm ell}^{(j)} \delta^{(1-r)/2}\right\} \quad \text{for } k \in \mathcal{P}_j^+$$
$$\mathbb{D}_{j,-}^{j+2} = \left\{\left|c_{j+2}^{(j)}\right| \leq m_{\rm adj}^{(j)} \left(C^{(j)}\delta\right)^{1/2}\right\}$$

and

$$\begin{aligned}\mathcal{N}_j^- = \Big\{ &\left(p_1^{(j)}, q_1^{(j)}, p_2^{(j)}, q_2^{(j)}\right) \in \mathbb{R}^4 : \\ &-C^{(j)}\delta\left(\ln(1/\delta) + m_{\rm hyp}^{(j)}\right) \leq p_1^{(j)} \leq -C^{(j)}\delta\left(\ln(1/\delta) - m_{\rm hyp}^{(j)}\right), \\ &q_1^{(j)} = \sigma,\ g_{\mathcal{I}_j}(p_2, q_2, \sigma, \delta) = 0,\ |p_2^{(j)}|, |q_2^{(j)}| \leq m_{\rm hyp}^{(j)}\left(C^{(j)}\delta\right)^{1/2}\Big\}.\end{aligned} \tag{4.31}$$

The function $g_{\mathcal{I}_j}(p_2, q_2, \sigma, \delta)$ is a smooth function defined in (6.5).

Remark 4.4. *Note that for this product-like sets the variable $p_1^{(j)}$ is selected negative. This is related to the fact that $\nu_{02} > 0$ (see Remark 4.2). The reason of the choice of the sign of $p_1^{(j)}$ will be clear in Section 5. In particular, see Remark 5.3.*

The domains $\mathcal{V}_j$ of the maps $\mathcal{B}^j$ will have $\mathcal{I}_j$-product-like structure as defined in Definition 4.3. Thus, we need to obtain the multi-parameter sets $\mathcal{I}_j$. They will be defined recursively. Recall that, to prove Theorem 3, we want to obtain an orbit which starts close to the periodic orbit $\mathbb{T}_3$. Thus, the recursively defined multi-parameter sets $\mathcal{I}_j$ will start with a set $\mathcal{I}_3$.

Definition 4.5. *Fix any constants $r, r' \in (0,1)$ satisfying $0 < r' < 1/2 - 2r$, $K > 0$ and small $\delta, \sigma > 0$. We say that a collection of multi-parameter sets $\{\mathcal{I}_j\}_{j=3,\dots,N-1}$ defined in (4.29) is (σ, δ, K)-recursive if for $j = 3, \dots, N-1$ the constants $C^{(j)}$ satisfy*

$$\begin{aligned} C^{(j)}/K \leq C^{(j+1)} &\leq KC^{(j)} \\ 0 < m_{\text{hyp}}^{(j+1)} &\leq m_{\text{hyp}}^{(j)} \end{aligned}$$

and all the other parameters should be strictly positive and are defined recursively as

$$\begin{aligned} M_{\text{ell},\pm}^{(j+1)} &= M_{\text{ell},\pm}^{(j)} + K\delta^{r'} \\ m_{\text{ell}}^{(j+1)} &= m_{\text{ell}}^{(j)} - K\delta^{r'} \\ M_{\text{adj},+}^{(j+1)} &= 2M_{\text{ell},+}^{(j)} + K\delta^{r'} \\ M_{\text{adj},-}^{(j+1)} &= KM_{\text{hyp}}^{(j)} \\ m_{\text{adj}}^{(j+1)} &= \frac{1}{2}m_{\text{ell}}^{(j)} - K\delta^{r'} \\ M_{\text{hyp}}^{(j+1)} &= KM_{\text{adj},+}^{(j)} \end{aligned}$$

The next Theorem defines recursively the product-like sets $\mathcal{V}_j$, so that condition (4.27) is satisfied.

Theorem 5 (Iterative Theorem)**.** *Fix large $\gamma > 0$, small $\sigma > 0$, and any constants $r, r' \in (0,1)$ satisfying $0 < r' < 1/2 - 2r$. Then, if we set $\delta = e^{-\gamma N}$, there exist strictly positive constants K and $C^{(3)}$ independent of N satisfying*

$$C^{(3)} \leq \delta^{-r} K^{-(N-2)}, \tag{4.32}$$

and a multi-parameter set $\mathcal{I}_3$ (as defined in (4.29)) with the following property: there exists a (σ, δ, K)-recursive collection of multi-parameter sets collection of multi-parameter sets $\{\mathcal{I}_j\}_{j=3,\dots,N-1}$ and $\mathcal{I}_j$-product-like sets $\mathcal{V}_j \subset \Sigma_j^{\text{in}}$ such that for each $j = 3, \dots, N-1$ we have

$$\mathcal{V}_{j+1} \subset \mathcal{B}^j(\mathcal{V}_j).$$

Moreover, the time spent to reach the section Σ_{j+1}^{in} can be bounded by

$$|T_{\mathcal{B}^j}| \leq K \ln(1/\delta)$$

for any $(p, q, c) \in \mathcal{V}_j$ and any $j = 3, \dots, N-2$.

Note that the condition

$$C^{(j)}/K < C^{(j+1)} < KC^{(j)}$$

implies

$$K^{-(j-2)}C^{(3)} \leq C^{(j+1)} \leq K^{j+2}C^{(3)}$$

Namely, at each saddle, the orbits we are studying may lie further from the heteroclinic orbit. Nevertheless, by the condition on δ from Theorem 3 and (4.32), these constant does not grow too much. Indeed,

$$\delta^r \leq C^{(j)} \leq \delta^{-r}, \tag{4.33}$$

where $r > 0$ can be taken as small as desired. We will use the bound (4.33) throughout the proof of Theorem 5.

Theorem 3 is a straightforward consequence of Theorem 5. In fact, we need more precise information than the one stated in Theorem 3. This more precise information will be used in the proof of Theorem 4. We state it in the following theorem. Theorem 3 is a straightforward consequence of it.

Theorem 3–bis *Assume that the conditions of Theorem 3 hold. Then, there exists an orbit $b(t)$ of equations* (3.12)*, constants $\mathbb{K} > 0$ and $\nu > 0$, independent of N and δ, and $T_0 > 0$ satisfying*

$$T_0 \leq \mathbb{K} N \ln(1/\delta),$$

such that

$$\begin{array}{lll} |b_3(0)| > 1 - \delta^\nu & & |b_{N-2}(T_0)| > 1 - \delta^\nu \\ |b_j(0)| < \delta^\nu \quad \text{for } j \neq 3 & \text{and} & |b_j(T_0)| < \delta^\nu \quad \text{for } j \neq N-2 \end{array}$$

Moreover, call $t_j \in [0, T_0]$ the times for which $b(t_j) \in \Sigma_j^{\text{in}}$, Then,

$$t_{j+1} - t_j \leq \mathbb{K} \ln(1/\delta)$$

and for any $t \in [t_j, t_{j+1}]$ and $k \neq j-1, j, j+1$,

$$|b_k(t)| \leq \delta^\nu.$$

Proof of Theorem 3–bis. It is enough to take as a initial condition b^0 a point in the set $\mathcal{V}_3 \subset \Sigma_3^{\text{in}}$ obtained in Theorem 5. Then, thanks to this theorem we know that there exists a time T_0 satisfying

$$T_0 \sim N \ln(1/\delta),$$

such that the corresponding orbit satisfies that $b(T_0) \in \mathcal{V}_{N-1} \subset \Sigma_{N-1}^{\text{in}}$. Note that in this section there are two components of b with size independent of δ. Nevertheless, from the proof of Theorem 5 in Section 6 it can be easily seen that if we shift the time interval $[0, T_0]$ to $[\rho \ln(1/\delta), \rho \ln(1/\delta) + T_0]$, for any $\rho < \sqrt{3}$, there exists $\nu > 0$ such that the orbit $b(t)$ satisfies the statements given in Theorem 3–bis. $\square$

4.3 Structure of the proof of the Iterative Theorem 5

To prove Theorem 5 we split it into two inductive lemmas. The first part analyzes the evolution of the trajectories close to the saddle j and the second one the travel along the heteroclinic orbit. Thus, we study $\mathcal{B}^j$ as a composition of two maps.

We consider an intermediate section transversal to the flow

$$\Sigma_j^{\text{out}} = \left\{ p_2^{(j)} = \sigma \right\}, \tag{4.34}$$

and then we consider two maps. First the local map

$$\mathcal{B}_{\text{loc}}^j : \mathcal{V}_j \subset \Sigma_j^{\text{in}} \longrightarrow \Sigma_j^{\text{out}}, \tag{4.35}$$

which studies the trajectories locally close to the saddle. Then, we consider a second map,

$$\mathcal{B}^j_{\mathrm{glob}} : \mathcal{U}^j \subset \Sigma^{\mathrm{out}}_j \longrightarrow \Sigma^{\mathrm{in}}_{j+1}, \tag{4.36}$$

which we call global map, that studies how the trajectories behave close to the heteroclinic orbit. Then, the map $\mathcal{B}^j$ considered in Theorem 5 is just $\mathcal{B}^j = \mathcal{B}^j_{\mathrm{glob}} \circ \mathcal{B}^j_{\mathrm{loc}}$.

Before we go into technicalities we write a table analogous to (4.4) of the properties of the local and global maps. The local map $\mathcal{B}^j_{\mathrm{loc}}$, projected onto hyperbolic variables, has the form

$$\begin{array}{lcl} p_1^{(j)} \sim C^{(j)}\ \delta\ \ln \dfrac{1}{\delta} & \longrightarrow & |p_1^{(j)}| \lesssim (C^{(j)}\delta)^{1/2} \\ q_1^{(j)} = \sigma & \longrightarrow & |q_1| \lesssim (C^{(j)}\delta)^{1/2} \\ |p_2^{(j)}| \lesssim (C^{(j)}\delta)^{1/2} & \longrightarrow & p_2^{(j)} = \sigma \\ |q_2^{(j)}| \lesssim (C^{(j)}\delta)^{1/2} & \longrightarrow & |q_2^{(j)}| \lesssim C^{(j)}\ \delta\ \ln \dfrac{1}{\delta}. \end{array} \tag{4.37}$$

The global map $\mathcal{B}^j_{\mathrm{glob}}$, projected onto hyperbolic variables of the corresponding saddles, has the form

$$\begin{array}{lcl} |p_1^{(j)}| \lesssim (C^{(j)}\delta)^{1/2} & \longrightarrow & |p_1^{(j+1)}| \lesssim C^{(j)}\delta\ \ln \dfrac{1}{\delta} \\ |q_1^{(j)}| \lesssim (C^{(j)}\delta)^{1/2} & \longrightarrow & q_1^{(j+1)} = \sigma \\ p_2^{(j)} = \sigma & \longrightarrow & |p_2^{(j+1)}| \lesssim (C^{(j)}\delta)^{1/2} \\ |q_2^{(j)}| \lesssim C^{(j)}\ \delta\ \ln \dfrac{1}{\delta} & \longrightarrow & |q_2^{(j+1)}| \lesssim (C^{(j)}\delta)^{1/2}. \end{array} \tag{4.38}$$

To compose the two maps we need that the set $\mathcal{U}^j$, introduced in (4.36), has a modified product-like structure. To define its properties, we consider the projection

$$\widetilde{\pi}\left(c_-^{(j)}, p_1^{(j)}, q_1^{(j)}, p_2^{(j)}, q_2^{(j)}, c_+^{(j)}\right) = \left(p_2^{(j)}, q_2^{(j)}, c_+^{(j)}\right).$$

Definition 4.6. *Fix constants $r \in (0,1)$, $\delta > 0$ and $\sigma > 0$ and define a multi-parameter set of positive constants*

$$\widetilde{\mathcal{I}}_j = \left\{\widetilde{C}^{(j)}, \widetilde{m}^{(j)}_{\mathrm{ell}}, \widetilde{M}^{(j)}_{\mathrm{ell},\pm},\ \widetilde{m}^{(j)}_{\mathrm{adj}},\ \widetilde{M}^{(j)}_{\mathrm{adj},\pm},\ \widetilde{m}^{(j)}_{\mathrm{hyp}},\ \widetilde{M}^{(j)}_{\mathrm{hyp}}\right\}.$$

Then, we say that a (non-empty) set $\mathcal{U} \subset \Sigma^{\mathrm{out}}_j$ has a $\widetilde{\mathcal{I}}_j$-product-like structure provided it satisfies the following two conditions:

C1

$$\mathcal{U} \subset \widetilde{\mathbb{D}}^1_j \times \ldots \times \widetilde{\mathbb{D}}^{j-2}_j \times \widetilde{\mathcal{N}}_{j,-} \times \widetilde{\mathbb{D}}^{j+2}_j \times \ldots \times \widetilde{\mathbb{D}}^N_j$$

where

$$\widetilde{\mathbb{D}}^k_j = \left\{\left|c_k^{(j)}\right| \le \widetilde{M}^{(j)}_{\mathrm{ell},\pm}\delta^{(1-r)/2}\right\} \quad \text{for } k \in \mathcal{P}_j^\pm$$

$$\widetilde{\mathbb{D}}^{j\pm2}_j \subset \left\{\left|c_{j\pm2}^{(j)}\right| \le \widetilde{M}^{(j)}_{\mathrm{adj},\pm}\left(\widetilde{C}^{(j)}\delta\right)^{1/2}\right\},$$

and

$$\begin{aligned} \widetilde{\mathcal{N}}_j^+ = \Big\{&(p_1^{(j)}, q_1^{(j)}, p_2^{(j)}, q_2^{(j)}) \in \mathbb{R}^4 : \left|p_1^{(j)}\right|, \left|q_1^{(j)}\right| \le \widetilde{M}^{(j)}_{\mathrm{hyp}}\left(\widetilde{C}^{(j)}\delta\right)^{1/2}, \\ &p_2^{(j)} = \sigma, -\widetilde{C}^{(j)}\,\delta\left(\ln(1/\delta) + \widetilde{M}^{(j)}_{\mathrm{hyp}}\right) \le q_2^{(j)} \le -\widetilde{C}^{(j)}\,\delta\left(\ln(1/\delta) - \widetilde{M}^{(j)}_{\mathrm{hyp}}\right)\Big\}, \end{aligned}$$

C2

$$\{\sigma\} \times \left[-\widetilde{C}^{(j)}\,\delta\,\left(\ln(1/\delta) - \widetilde{m}_{\mathrm{hyp}}^{(j)}\right), -\widetilde{C}^{(j)}\,\delta\,\left(\ln(1/\delta) + \widetilde{m}_{\mathrm{hyp}}^{(j)}\right)\right] \times \mathbb{D}_{j,-}^{j+2} \times \ldots \times \mathbb{D}_{j,-}^{N} \subset \widetilde{\pi}(\mathcal{U})$$

where

$$\begin{aligned}\mathbb{D}_{j,-}^{k} &= \left\{\left|c_k^{(j)}\right| \leq \widetilde{m}_{\mathrm{ell}}^{(j)}\delta^{(1-r)/2}\right\} \quad \textit{for } k \in \mathcal{P}_j^+\\ \mathbb{D}_{j,-}^{j+2} &= \left\{\left|c_{j+2}^{(j)}\right| \leq \widetilde{m}_{\mathrm{adj}}^{(j)}\left(C^{(j)}\delta\right)^{1/2}\right\}.\end{aligned}$$

With this definition, we can state the following two lemmas. Combining these two lemmas we deduce Theorem 5.

Lemma 4.7. *Fix any natural j with $3 \leq j \leq N-2$, constants $r, r' \in (0,1)$ satisfying $0 < r' < 1/2 - 2r$ and $\sigma > 0$ small enough. Take $\delta = e^{-\gamma N}$, $\gamma = \gamma(\sigma) \gg 1$, depending on σ, and consider a parameter set $\mathcal{I}_j$ with $M_{\mathrm{hyp}}^{(j)} \geq 1$ and a $\mathcal{I}_j$-product-like set $\mathcal{V}_j \subset \Sigma_j^{\mathrm{in}}$. Then, f or N big enough, there exists:*

- *A constant $K > 0$ independent of N and j but which might depend on σ.*
- *A parameter set $\widetilde{\mathcal{I}}_j$ whose constants satisfy*

$$\begin{aligned}C^{(j)}/2 \leq \widetilde{C}^{(j)} &\leq 2C^{(j)}\\ 0 < \widetilde{m}_{\mathrm{hyp}}^{(j)} &\leq m_{\mathrm{hyp}}^{(j)}\end{aligned}$$

and

$$\begin{aligned}\widetilde{M}_{\mathrm{hyp}}^{(j)} &= K\\ \widetilde{M}_{\mathrm{ell},\pm}^{(j)} &= M_{\mathrm{ell},\pm}^{(j)} + K\delta^{r'}\\ \widetilde{m}_{\mathrm{ell}}^{(j)} &= m_{\mathrm{ell}}^{(j)} - K\delta^{r'}\\ \widetilde{M}_{\mathrm{adj},\pm}^{(j)} &= M_{\mathrm{adj},\pm}^{(j)}(1+4\sigma)\\ \widetilde{m}_{\mathrm{adj}}^{(j)} &= m_{\mathrm{adj}}^{(j)}\,(1-4\sigma),\end{aligned}$$

- *A $\widetilde{\mathcal{I}}_j$-product-like set $\mathcal{U}_j$ for which the map $\mathcal{B}_{\mathrm{loc}}^{j}$ satisfies*

$$\mathcal{U}_j \subset \mathcal{B}_{\mathrm{loc}}^{j}\left(\mathcal{V}_j\right). \tag{4.39}$$

Moreover, the time to reach the section Σ_j^{out} can be bounded as

$$\left|T_{\mathcal{B}_{\mathrm{loc}}^{j}}\right| \leq K\ln(1/\delta).$$

The proof of this lemma is the *most delicate part* in the proof of the Iterative Theorem 5, since we are passing close to a hyperbolic fixed point, which implies big deviations. It is split in several parts in the forthcoming sections to simplify the exposition.

First, in Section 5, we set the elliptic modes c to zero, and we study the saddle map associated to the corresponding system. We call to this system *Hyperbolic Toy Model.* It has two degrees of freedom. The saddle is resonant since both stable eigenvalues coincide (see (4.14)) and therefore, this Hyperbolic Toy Model is not well approximated by its linearization around the saddle. This fact complicates the proof of Lemma 4.7 and it has been exemplified with a simplified model in Section 2.2. To overcome this problem, we consider the techniques developed by Shilnikov [Šil67], which allow us to consider a good nonlinear first order of the Hyperbolic Toy Model which gives a very precise control of the behavior of the Hyperbolic Toy Model while traveling close to the saddle.

Then, in Section 6 we use the results obtained for the Hyperbolic Toy Model to deal with the full system and prove Lemma 4.7. To prove the lemma we take advantage of the fact that, since we take the elliptic modes rather small, at first order they are just rotating and therefore their modulus barely change. This implies that at first order, the coupling between the elliptic and the hyperbolic modes is very weak and thus, using the results of the Hyperbolic Toy Model with some additional analysis of the elliptic modes, one can prove Lemma 4.7.

Now we state the iterative lemma for the global maps $\mathcal{B}^j_{\mathrm{glob}}$.

Lemma 4.8. *Fix any natural j with $3 \le j \le N-2$, constants $r, r' \in (0,1)$ satisfying $0 < r' < 1/2 - 2r$ and $\sigma > 0$ small enough. Take $\delta = e^{-\gamma N}$, $\gamma = \gamma(\sigma) \gg 1$, depending on σ, and consider a parameter set $\widetilde{\mathcal{I}}_j$ and a $\widetilde{\mathcal{I}}_j$-product-like set $\mathcal{U}_j \subset \Sigma_j^{\mathrm{out}}$. Then, for N large enough, there exists:*

- *A constant $\widetilde{K}$ depending on σ, but independent of N and j.*
- *A parameter set $\mathcal{I}_{j+1}$ whose constants satisfy*

$$\widetilde{C}^{(j)}/\widetilde{K} \le C^{(j+1)} \le \widetilde{K}\widetilde{C}^{(j)}$$
$$0 < m_{\mathrm{hyp}}^{(j+1)} \le \widetilde{m}_{\mathrm{hyp}}^{(j)}$$

  and

$$\begin{aligned}
M_{\mathrm{ell},-}^{(j+1)} &= \max\left\{\widetilde{M}_{\mathrm{ell},-}^{(j)} + \widetilde{K}\delta^{r'}, \widetilde{K}\widetilde{M}_{\mathrm{adj},-}^{(j)}\right\}\\
M_{\mathrm{ell},+}^{(j+1)} &= \widetilde{M}_{\mathrm{ell},+}^{(j)} + \widetilde{K}\delta^{r'}\\
m_{\mathrm{ell}}^{(j+1)} &= \widetilde{m}_{\mathrm{ell}}^{(j)} - \widetilde{K}\delta^{r'}\\
M_{\mathrm{adj},+}^{(j+1)} &= \widetilde{M}_{\mathrm{ell},+}^{(j)} + \widetilde{K}\delta^{r'}\\
M_{\mathrm{adj},-}^{(j+1)} &= \widetilde{K}\widetilde{M}_{\mathrm{hyp}}^{(j)}\\
m_{\mathrm{adj}}^{(j+1)} &= \widetilde{m}_{\mathrm{ell}}^{(j)} + \widetilde{K}\delta^{r'}\\
M_{\mathrm{hyp}}^{(j+1)} &= \max\left\{\widetilde{K}\widetilde{M}_{\mathrm{adj},+}^{(j)}, \widetilde{K}\right\}
\end{aligned}$$

- *A $\mathcal{I}_{j+1}$-product-like set $\mathcal{V}_{j+1} \subset \Sigma_{j+1}^{\mathrm{in}}$ for which the map $\mathcal{B}^j_{\mathrm{glob}}$ satisfies*

$$\mathcal{V}_{j+1} \subset \mathcal{B}^j_{\mathrm{glob}}(\mathcal{U}_j). \tag{4.40}$$

Moreover, the time spent to reach the section $\Sigma_{j+1}^{\mathrm{in}}$ can be bounded as

$$\left|T_{\mathcal{B}^j_{\mathrm{glob}}}\right| \le \widetilde{K}.$$

The proofs of this lemma is postponed to Section 7.

Now it only remains to deduce from Lemmas 4.7 and 4.8 the Iterative Theorem 5.

Proof of Theorem 5. We choose the multiindex $\mathcal{I}_3$ so that we can apply iteratively the Lemmas 4.7 and 4.8. Indeed, from the recursive formulas in Lemma 4.7 and 4.8 it is clear that it is enough to chose a parameter set $\mathcal{I}_3$ satisfying

$$1 < M^{(3)}_{\text{ell},+} \ll M^{(3)}_{\text{adj},+} \ll M^{(3)}_{\text{hyp}} \ll M^{(3)}_{\text{adj},-} \ll M^{(3)}_{\text{ell},-}$$

and

$$0 < m^{(3)}_{\text{ell}} < 3m^{(3)}_{\text{adj}}.$$

From the choice of the constants in $\mathcal{I}_3$ and the recursion formulas in Lemmas 4.7 and 4.8, we have that $M^{(j)}_{\text{hyp}} \geq 1$ for any $j = 3, \dots N-1$. This fact along with conditions (4.39) and (4.40), allow us to apply Lemmas 4.7 and 4.8 iteratively so that we obtain the (δ, σ, K)-recursive collection of multi-parameter sets $\{\mathcal{I}_j\}_{j=3,\dots,N-1}$ and the $\mathcal{I}_j$-product-like sets $\mathcal{V}_j \subset \Sigma^{\text{in}}_j$. In particular, note that the recursion formulas stated in Theorem 5 can be easily deduced from the recursion formulas given in Lemmas 4.7 and 4.8 and the choice of $\mathcal{I}_3$.

Finally, we bound the time

$$|T_{\mathcal{B}^j}| \leq \left|T_{\mathcal{B}^j_{\text{loc}}}\right| + \left|T_{\mathcal{B}^j_{\text{glob}}}\right| \leq (K + \widetilde{K}) \ln(1/\delta).$$

This completes the proof of Theorem 5. $\square$

5 The Hyperbolic Toy Model

In this section we set the elliptic modes to zero, namely, we deal with the system

$$\begin{aligned}
\dot{p}_1 &= \sqrt{3}p_1 + \mathcal{Z}_{\text{hyp},p_1} \\
\dot{q}_1 &= -\sqrt{3}q_1 + \mathcal{Z}_{\text{hyp},q_1} \\
\dot{p}_2 &= \sqrt{3}p_2 + \mathcal{Z}_{\text{hyp},p_2} \\
\dot{q}_2 &= -\sqrt{3}q_2 + \mathcal{Z}_{\text{hyp},q_2},
\end{aligned} \tag{5.1}$$

where the functions $\mathcal{Z}_{\text{hyp},*}$ are defined in (4.15), (4.16), (4.17) and (4.18).

We start by setting some notation. We call

$$z = (x_1, y_1, x_2, y_2)$$

the new set of coordinates, whose components are also denoted by $z_i = (x_i, y_i)$. We also use the notation $x = (x_1, x_2)$ and $y = (y_1, y_2)$.

Moreover, we call K to any positive constant independent of δ, N, j, and σ and we call K_σ to any positive constant depending on σ, but independent of δ, N and j Analogously, we say that $a = \mathcal{O}(b)$ if $|a| \leq K|b|$ and that $a = \mathcal{O}_\sigma(b)$ if $|a| \leq K_\sigma|b|$. We will also use all these notations in Section 6 and Section 7.

The first step is to perform a resonant $\mathcal{C}^k$ normal form in a neighborhood of size σ of the saddle. Note that we do not need much regularity for the normal form since all our study will be done in the $\mathcal{C}^0$ norm. It turns out it is enough to consider a $\mathcal{C}^1$ normal form. Before we state our next claim about the normal form we formulate a well known result of Bronstein-Kopanskii

[BK92] about finitely smooth normal forms of vector fields near a critical point. We are unable to use classical results about linearizability, because our saddle is *resonant*.

The main result of Bronstein-Kopanskii [BK92] is that near a saddle point a vector field can be transformed into a polynomial one by a finitely smooth change of coordinates with only certain (resonant) monomials present. For convenience of the reader we use notations of this paper.

5.1 Finitely smooth polynomial normal forms of vector fields in near a saddle point

Let $\dot{x} = F(x)$ be a vector with the origin being a critical point, i.e. $F(0) = 0,\ x \in \mathbb{R}^d$ for some $d \in \mathbb{Z}_+$. Assume that F is C^K for some positive integer $K \in \mathbb{Z}_+$, i.e. F has all partial derivatives of order up to K uniformly bounded. Denote the linearization of F at 0 by $A := DF(0)$ and $f(x) = F(x) - A(x)$. Then, the equation becomes

$$\dot{x} = Ax + f(x),\ f(0) = 0,\ Df(0) = 0.$$

Let $\nu_1, \dots, \nu_d$ denote the eigenvalues of A and $\theta_1, \dots, \theta_n$ be all distinct numbers contained in the set $\{\operatorname{Re}\nu_i :\ i = 1, \dots, d\}$. Assume that none of θ_i's is zero or, in other words, the rest point being hyperbolic.

The space $\mathbb{R}^d$ can be represented as a direct sum of A-invariant subspaces $E_1, \dots, E_n$ such that the eigenvalues of the operator $A|_{E_i}$ satisfy the condition $\operatorname{Re}\nu_i = \theta_i$.

Theorem 6. *[BK92] Let k be positive integer. Assume that the vector field $\dot{x} = F(x)$ is of class C^K, $x = 0$ is a hyperbolic saddle point and $A = DF(0)$. If $K \geq Q(k)$ for some computable function $Q(\cdot)$, then, for some positive integer N, this vector field near the point $x = 0$ can be reduced by a transformation $y = \Phi(x), \Phi \in C^k$, to the polynomial resonant normal form*

$$\dot{y} = Ay + \sum_{|\tau|=2}^{N} p_\tau y^\tau,$$

where $\tau \in \mathbb{Z}_+^d$ and p_τ denotes a multi-homogeneous polynomial $p_\tau(E_1, \dots, E_n; E_1 \oplus \dots \oplus E_n)$, $p_\tau = (p_\tau^1, \dots, p_\tau^d)$ and $p_\tau^i \neq 0$ implies $\nu_i = \tau^1\nu_1 + \dots + \tau^d\nu_d$ (by the resonant condition).

In Theorem 3 [BK92] the authors give an upper bound on N. In our case $d = 4,\ n = 2,\ k = 1$. A direct application of this Theorem is the following

Lemma 5.1. *There exists a $\mathcal{C}^1$ change of coordinates*

$$(p_1, q_1, p_2, q_2) = \Psi_{\text{hyp}}(x_1, y_1, x_2, y_2) = (x_1, y_1, x_2, y_2) + \widetilde{\Psi}_{\text{hyp}}(x_1, y_1, x_2, y_2)$$

which transforms the vector field (5.1) *into the vector field*

$$\mathcal{X}_{\text{hyp}}(z) = Dz + R_{\text{hyp}}, \tag{5.2}$$

where D is the diagonal matrix $D = \operatorname{diag}(\sqrt{3}, -\sqrt{3}, \sqrt{3}, -\sqrt{3})$ and R_{hyp} is a polynomial, which only contains resonant monomials [6]*. It can be split as*

$$R_{\text{hyp}} = R^0_{\text{hyp}} + R^1_{\text{hyp}}, \tag{5.3}$$

[6] If the reader is interested in bounding the degree of the polynomial see Theorem 3 in page 169 of [BK92] for an estimate. Nevertheless, in the present paper, we just use that R_{hyp} is a polynomial and thus has some finite degree

where $R^0_{\rm hyp}$ is the first order, which is given by

$$R^0_{\rm hyp}(z) = \begin{pmatrix} R^0_{{\rm hyp},x_1}(z) \\ R^0_{{\rm hyp},y_1}(z) \\ R^0_{{\rm hyp},x_2}(z) \\ R^0_{{\rm hyp},y_2}(z) \end{pmatrix} = \begin{pmatrix} 2\nu_2 x_1^2 y_1 + 2\nu_{02} y_1 x_2^2 + \nu_{11} x_1 x_2 y_2 \\ -2\nu_2 x_1 y_1^2 - 2\nu_{20} x_1 y_2^2 - \nu_{11} y_1 x_2 y_2 \\ 2\nu_2 y_2 x_2^2 + 2\nu_{20} x_1^2 y_2 + \nu_{11} x_1 y_1 x_2 \\ -2\nu_2 x_2 y_2^2 - \nu_{02} y_1^2 x_2 - \nu_{11} x_1 y_1 y_2 \end{pmatrix}, \tag{5.4}$$

and $R^1_{\rm hyp}$ is the remainder and satisfies

$$R^1_{{\rm hyp},x_i} = \mathcal{O}\left(x^3 y^2\right) \quad \textit{and} \quad R^1_{{\rm hyp},y_i} = \mathcal{O}\left(x^2 y^3\right). \tag{5.5}$$

Moreover, the function $\widetilde{\Psi}_{\rm hyp} = (\widetilde{\Psi}_{{\rm hyp},x_1}, \widetilde{\Psi}_{{\rm hyp},y_1}, \widetilde{\Psi}_{{\rm hyp},x_2}, \widetilde{\Psi}_{{\rm hyp},y_2})$ satisfies

$$\begin{aligned} \widetilde{\Psi}_{{\rm hyp},x_1}(z) &= \mathcal{O}\left(x_1^3, x_1 y_1, x_1(x_2^2 + y_2^2), y_1 y_2 (x_2 + y_2)\right) \\ \widetilde{\Psi}_{{\rm hyp},y_1}(z) &= \mathcal{O}\left(y_1^3, x_1 y_1, y_1(x_2^2 + y_2^2), x_1 x_2 (x_2 + y_2)\right) \\ \widetilde{\Psi}_{{\rm hyp},x_2}(z) &= \mathcal{O}\left(x_2^3, x_2 y_2, x_2(x_1^2 + y_1^2), y_1 y_2 (x_1 + y_1)\right) \\ \widetilde{\Psi}_{{\rm hyp},y_2}(z) &= \mathcal{O}\left(y_2^3, x_2 y_2, y_2(x_1^2 + y_1^2), x_1 x_2 (x_1 + y_1)\right). \end{aligned}$$

5.2 The local map for the Hyperbolic Toy Model in the normal form variables

Recall that our goal in this step of the proof is to study the evolution of points with initial conditions inside of a certain set near the section $\Sigma_j^{\rm in}$. More specifically, in formulas (4.30) and (4.31) we define sets $\mathcal{N}_j^- \subset \mathcal{N}_j^+$. We set elliptic modes $c = 0$ and shall study the set $\mathcal{N}_j'$ satisfying

$$\mathcal{N}_j^- \cap \{c = 0\} \subset \mathcal{N}_j' \subset \mathcal{N}_j^+ \cap \{c = 0\}.$$

Since the analysis is done in normal coordinates $\Psi_{\rm hyp} : (x, y) \to (p, q)$, we study the a set $\widehat{\mathcal{N}}_j$ such that $\Psi_{\rm hyp}^{-1}(\mathcal{N}_j') \subset \widehat{\mathcal{N}}_j$. To define this set we need to fix several parameters and define several objects.

Let $C^{(j)}$'s be the constant from Lemma 4.7. Recall that in Definition 4.5 we define a (σ, δ, K)-recursive multiparameter set $\mathcal{I}_j$. Its description includes parameters $M^{(j)}_{\rm hyp}$ used below. The parameter K depends on σ and we keep this dependence in the notation: K_σ. Denote the inverse of the map Ψ from Lemma 5.1, by

$$\Upsilon := {\rm Id} + \widetilde{\Upsilon} := \Psi_{\rm hyp}^{-1} =: {\rm Id} + (\widetilde{\Upsilon}_{x_1}, \widetilde{\Upsilon}_{y_1}, \widetilde{\Upsilon}_{x_2}, \widetilde{\Upsilon}_{y_2}).$$

Define

$$\widehat{C}^{(j)} := C^{(j)} \left(1 + \partial_{x_1} \widetilde{\Upsilon}_{x_1}(0, \sigma, 0, 0)\right). \tag{5.6}$$

Notice that $\widehat{C}^{(j)} = C^{(j)} \left(1 + \mathcal{O}(\sigma)\right)$. Define $f_1(\sigma)$ by

$$f_1(\sigma) = \Upsilon_{y_1}(0, \sigma, 0, 0). \tag{5.7}$$

Observe that it satisfies $f_1(\sigma) = \sigma + \mathcal{O}(\sigma^3)$ and the section $\{y_1 = f_1(\sigma)\}$ approximates the image of the section $\Upsilon(\Sigma_j^{\rm in})$. Now we can define the set of points whose evolution under the local map we shall analyze

$$\begin{aligned} \widehat{\mathcal{N}}_j = \Big\{ &|x_1 + \widehat{C}^{(j)} \delta \left(\ln(1/\delta)\right| \le \widehat{C}^{(j)} \delta K_\sigma, \quad |x_2 - x_2^*| \le 2 M^{(j)}_{\rm hyp} \frac{\left(\widehat{C}^{(j)} \delta\right)^{1/2}}{\ln(1/\delta)}, \\ &|y_1 - f_1(\sigma)| \le K_\sigma \widehat{C}^{(j)} \delta \ln(1/\delta), \qquad |y_2| \le 2 M^{(j)}_{\rm hyp} \left(\widehat{C}^{(j)} \delta\right)^{1/2} \Big\}, \end{aligned} \tag{5.8}$$

where the constant x_2^* will be defined later in this section. It turns out a proper choice of x_2^* leads to a cancelation in the evolution of the x_1 coordinate (described in Section 2.2 for the simplified model). This cancelation is crucial to obtain good estimates for the map $\mathcal{B}^j_{\mathrm{loc}}$.

We also define the function $f_2(\sigma)$ as

$$f_2(\sigma) = \Upsilon_{x_2}(0, 0, \sigma, 0). \tag{5.9}$$

By analogy with $f_1(\sigma)$ notice that the section $\{x_2 = f_2(\sigma)\}$ approximates the image of the section $\Upsilon(\Sigma_j^{\mathrm{out}})$ with $\Sigma_j^{\mathrm{out}} = \{p_2 = \sigma\}$. Later we need to compute an approximate transition time $T_j(x_2)$ from near $\Upsilon(\Sigma_j^{\mathrm{in}})$ to $\Upsilon(\Sigma_j^{\mathrm{out}})$. We use f_2 to do that. Notice that the x_2 coordinate behaves almost linearly as

$$x_2 \sim x_2^0 e^{\sqrt{3}t}.$$

Therefore, for an orbit to reach $\{x_2 = f_2(\sigma)\}$ it takes an approximate time

$$T_j\left(x_2^0\right) = \frac{1}{\sqrt{3}} \ln\left(\frac{f_2(\sigma)}{x_2^0}\right). \tag{5.10}$$

Note that this time is defined for any $x_2^0 > 0$. We will see that the x_2^0 coordinate behaves as $x_2^0 \sim (\widehat{C}^{(j)}\delta)^{1/2}$ and, therefore, T_j behaves as

$$T_j \sim \ln \frac{1}{\widehat{C}^{(j)}\delta}.$$

Even if x_2 behaves approximately as for a linear system, this is not the case for the other variables, as we have explained in Section 2.2 with a simplified model. Indeed, if one first considers the linear part of the vector field (5.1), omiting the dependence on $\widehat{C}^{(j)}$, the transition map sends points

$$(x_1, y_1, x_2, y_2) \sim \left(\mathcal{O}(\delta\ln(1/\delta)), \mathcal{O}(\sigma), \mathcal{O}\left(\delta^{1/2}\right), \mathcal{O}\left(\delta^{1/2}\right)\right)$$

to

$$(x_1, y_1, x_2, y_2) \sim \left(\mathcal{O}\left(\delta^{1/2}\ln(1/\delta)\right), \mathcal{O}\left(\delta^{1/2}\right), \mathcal{O}(\sigma), \mathcal{O}(\delta)\right).$$

However, the resonance implies a certain deviation from the heteroclinic orbits. Indeed, one can see that tipically, the image point is of the form

$$(x_1, y_1, x_2, y_2) \sim \left(\mathcal{O}\left(\delta^{1/2}\ln(1/\delta)\right), \mathcal{O}\left(\delta^{1/2}\right), \mathcal{O}(\sigma), \mathcal{O}(\delta\ln(1/\delta))\right).$$

This apparently small deviation, after undoing the normal form, would imply a considerably big deviation from the heteroclinic orbit and would lead to very bad estimates. Nevertheless, if one chooses carefully x_2 in terms of x_1 and y_1, one can obtain a cancelation that leads to an image point of the form

$$(x_1, y_1, x_2, y_2) \sim \left(\mathcal{O}\left(\delta^{1/2}\right), \mathcal{O}\left(\delta^{1/2}\right), \mathcal{O}(\sigma), \mathcal{O}(\delta\ln(1/\delta))\right).$$

Since the points we are dealing with belong to the set $\widetilde{\mathcal{N}}_j$ defined in (5.8), this cancellation boils down to choosing a suitable constant x_2^*. Next lemma shows that a particular choice of x_2^* leads to a cancellation that allow us to obtain good estimates for the saddle map in spite of the resonance. The choice we do is essentially the same as the one choosen in Section 2.2 for the simplified model that has been considered in that section.

Lemma 5.2. *Let us consider the flow Φ_t^{hyp} associated to* (5.2) *and a point $z^0 \in \widehat{\mathcal{N}}_j$. Then, if we choose x_2^* as the unique positive solution of*

$$(x_2^*)^2\, T_j(x_2^*) = \frac{\widehat{C}^{(j)}\, \delta\, \ln(1/\delta)}{2\, \nu_{02}\, f_1(\sigma)} \tag{5.11}$$

and we take δ and σ small enough, the point

$$z^f = \Phi_{T_j}^{\mathrm{hyp}}(z^0),$$

where $T_j = T_j(x_2^0)$ is the time defined in (5.10), *satisfies*

$$\begin{aligned}
|x_1^f| &\leq K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2}\\
|y_1^f| &\leq K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2}\\
|x_2^f - f_2(\sigma)| &\leq K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2} \ln^2(1/\delta)\\
\left| y_2^f + \frac{f_1(\sigma)}{f_2(\sigma)}\widehat{C}^{(j)}\delta \ln(1/\delta)\right| &\leq K_\sigma \widehat{C}^{(j)}\delta.
\end{aligned}$$

Remark 5.3. *The particular choice of x_2^* being a solution* (5.11) *will ensure a cancellation. This cancellation is crucial to obtain good estimates for the local map.*

Equation (5.11) *has real solutions because $\nu_{02} > 0$ (see Remark 4.2) and $x_1 < 0$ (and $p_1 < 0$ in the original variables, see Remark 4.4). Indeed, if $x_1 > 0$ and $x_1 \sim \widehat{C}^{(j)}\delta \ln(1/\delta)$ we have*

$$(x_2^*)^2\, T_j(x_2^*) = -\frac{\widehat{C}^{(j)}\, \delta\, \ln(1/\delta)}{2\, \nu_{02}\, f_1(\sigma)}.$$

If there is no solution to this equation, we cannot attain the desired cancellation.

Let us point out that taking into account the estimates for the points in $\widehat{\mathcal{N}}^{(j)}$, the definition of T_j in (5.10) and condition (4.33), one can deduce that condition (5.11) implies

$$|x_2^*| \leq K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2} \leq K_\sigma \delta^{(1-r)/2}.$$

and then,

$$T_j(x_2^0) \leq K_\sigma \ln(1/\delta). \tag{5.12}$$

We use this estimate throughout the proof of Lemma 5.2. Note also that for the modes (x_1^f, y_1^f) we just need upper bounds, since after the passage of the saddle j, the associated mode will become elliptic and therefore we will not need accurate estimates anymore.

Proof of Lemma 5.2. We prove the lemma using a fixed point argument. We look for a contractive operator using the variation of constants formula. Namely, we perform the change of coordinates

$$x_i = e^{\sqrt{3}t}u_i,\ y_i = e^{-\sqrt{3}t}v_i \tag{5.13}$$

and then we obtain the integral equations

$$\begin{aligned}
u_i &= x_i^0 + \int_0^T e^{-\sqrt{3}t} R_{\mathrm{hyp},x_i}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}\right) dt\\
v_i &= y_i^0 + \int_0^T e^{\sqrt{3}t} R_{\mathrm{hyp},y_i}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}\right) dt.
\end{aligned} \tag{5.14}$$

In the linear case u_i's and v_i's are fixed. We use these variables to find a fixed point argument. We define the contractive operator in two steps. This approach is inspired by Shilnikov [Šil67].

First we define an auxiliary (non-contractive) operator we follows

$$\mathcal{F}_{\rm hyp} = (\mathcal{F}_{{\rm hyp},u_1}, \mathcal{F}_{{\rm hyp},v_1}, \mathcal{F}_{{\rm hyp},u_2}, \mathcal{F}_{{\rm hyp},v_2})$$

as

$$\begin{aligned} \mathcal{F}_{{\rm hyp},u_i}(u,v) &= x_i^0 + \int_0^T e^{-\sqrt{3}t} R_{{\rm hyp},x_i}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}\right) dt \\ \mathcal{F}_{{\rm hyp},v_i}(u,v) &= y_i^0 + \int_0^T e^{\sqrt{3}t} R_{{\rm hyp},y_i}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}\right) dt. \end{aligned} \tag{5.15}$$

One can easily see that in the u_1 and v_2 components the main terms are *not* given by the initial condition *but by the integral terms.* This indicates that the dynamics near the saddle is *not* well approximated by the linearized dynamics and the operator is not contractive.

Following ideas from Shilnikov [Šil67], we modify slightly two of the components of $\mathcal{F}_{\rm hyp}$ and obtain a contractive operator. We define a new operator

$$\widetilde{\mathcal{F}}_{\rm hyp} = (\widetilde{\mathcal{F}}_{{\rm hyp},u_1}, \widetilde{\mathcal{F}}_{{\rm hyp},v_1}, \widetilde{\mathcal{F}}_{{\rm hyp},u_2}, \widetilde{\mathcal{F}}_{{\rm hyp},v_2})$$

as

$$\begin{aligned} \widetilde{\mathcal{F}}_{{\rm hyp},u_1}(u_1,v_1,u_2,v_2) &= \mathcal{F}_{{\rm hyp},u_1}(u_1, \mathcal{F}_{{\rm hyp},v_1}(u_1,v_1,u_2,v_2), \mathcal{F}_{{\rm hyp},u_2}(u_1,v_1,u_2,v_2), v_2) \\ \widetilde{\mathcal{F}}_{{\rm hyp},v_1}(u_1,v_1,u_2,v_2) &= \mathcal{F}_{{\rm hyp},v_1}(u_1,v_1,u_2,v_2) \\ \widetilde{\mathcal{F}}_{{\rm hyp},u_2}(u_1,v_1,u_2,v_2) &= \mathcal{F}_{{\rm hyp},u_2}(u_1,v_1,u_2,v_2) \\ \widetilde{\mathcal{F}}_{{\rm hyp},v_2}(u_1,v_1,u_2,v_2) &= \mathcal{F}_{{\rm hyp},v_2}(u_1, \mathcal{F}_{{\rm hyp},v_1}(u_1,v_1,u_2,v_2), \mathcal{F}_{{\rm hyp},u_2}(u_1,v_1,u_2,v_2), v_2) \end{aligned} \tag{5.16}$$

Note that the fixed points of these operators are exactly the same as the fixed points of $\mathcal{F}_{\rm hyp}$. Thus, the fixed points of the operator $\widetilde{\mathcal{F}}_{\rm hyp}$ are solutions of equation (5.14).

It turns out the operator $\widetilde{\mathcal{F}}_{\rm hyp}$ is contractive in a suitable Banach space. We define the following weighted norms. To fix notation, we denote by $\|\cdot\|_\infty$ the standard supremum norm. Then define

$$\begin{aligned} \|h\|_{{\rm hyp},u_1} &= \sup_{t\in[0,T_j]} \left| \left(-\widehat{C}^{(j)}\delta\ln(1/\delta) + 2\nu_{02} f_1(\sigma)\,(x_2^*)^2\, t + \widehat{C}^{(j)}\delta\right)^{-1} h(t)\right| \\ \|h\|_{{\rm hyp},v_1} &= f_1(\sigma)^{-1}\|h\|_\infty \\ \|h\|_{{\rm hyp},u_2} &= (x_2^*)^{-1}\,\|h\|_\infty \\ \|h\|_{{\rm hyp},v_2} &= \left(\left(y_1^0\right)^2 x_2^0 T_j\right)^{-1} \|h\|_\infty \end{aligned} \tag{5.17}$$

and the norm

$$\|(u,v)\|_* = \sup_{i=1,2} \left\{ \|u_i\|_{{\rm hyp},u_i}, \|v_i\|_{{\rm hyp},v_i} \right\}. \tag{5.18}$$

This gives rise to the following Banach space

$$\mathcal{Y}_{\rm hyp} = \left\{ (u,v) : [0,T] \to \mathbb{R}^4; \|(u,v)\|_* < \infty \right\}.$$

The contractivity of $\widetilde{\mathcal{F}}_{\rm hyp}$ is a consequence of the following two auxiliary propositions.

Proposition 5.4. *Assume* (5.11)*, then there exists a constant* $\kappa_0 > 0$ *independent of* σ*,* δ *and* j *such that for* δ *and* σ *small enough, the operator* $\widetilde{\mathcal{F}}_{\rm hyp}$ *satisfies*

$$\|\widetilde{\mathcal{F}}(0)\|_* \leq \kappa_0.$$

Proposition 5.5. *Consider* $w, w' \in B(2\kappa_0) \subset \mathcal{Y}_{\rm hyp}$ *and let us assume* (5.11)*, then taking* $\delta \ll \sigma$*, the operator* $\widetilde{\mathcal{F}}_{\rm hyp}$ *satisfies*

$$\|\widetilde{\mathcal{F}}_{\rm hyp}(w) - \widetilde{\mathcal{F}}_{\rm hyp}(w')\|_* \leq K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2} \ln^2(1/\delta)\|w - w'\|_*.$$

These two propositions show that $\widetilde{\mathcal{F}}_{\rm hyp}$ is contractive from $B(2\kappa_0) \subset \mathcal{Y}_{\rm hyp}$ to itself. Moreover, using them we can deduce accurate estimates for the image point. We prove here Proposition 5.4. The proof of Proposition 5.5 is deferred to the end of the section.

Proof of Proposition 5.4. We bound each mode separately. For $\widetilde{\mathcal{F}}_{{\rm hyp},v_1}$ and $\widetilde{\mathcal{F}}_{{\rm hyp},u_2}$, we have that

$$\widetilde{\mathcal{F}}_{{\rm hyp},v_1}(0) = y_1^0 \quad \text{and} \quad \widetilde{\mathcal{F}}_{{\rm hyp},u_2}(0) = x_2^0$$

and therefore, they satisfy the desired bounds. Now we bound the first iteration for u_1. Here we use the particular choice of x_2^0 in terms of (x_1^0, y_1^0) done in (5.11) to obtain the desired cancellations (see Remark 5.3). Indeed, taking into account the properties of $R_{{\rm hyp},x_1}$ given in Lemma 5.1, the first iteration is just

$$\begin{aligned}\widetilde{\mathcal{F}}_{{\rm hyp},u_1}(0)(t) =& x_1^0 + \int_0^t \left(2\nu_{02} y_1^0 (x_2^0)^2 + \mathcal{O}((y_1^0)^2(x_2^0)^3)\right) dt\\ & x_1^0 + 2\nu_{02} y_1^0 (x_2^0)^2 t + \mathcal{O}\left((y_1^0)^2(x_2^0)^3\right).\end{aligned}$$

Therefore, taking into account that $z^0 \in \widehat{\mathcal{N}}_j$ (see (5.8)) and also (5.12), we have that

$$\widetilde{\mathcal{F}}_{{\rm hyp},u_1}(0)(t) = -\widehat{C}^{(j)}\delta\ln(1/\delta) + 2\nu_{02} f_1(\sigma)(x_2^*)^2 t + \mathcal{O}\left(\widehat{C}^{(j)}\delta\right).$$

Thus, applying the norm given in (5.17), we have that there exists a constant $\kappa_0 > 0$ such that

$$\left\|\widetilde{\mathcal{F}}_{{\rm hyp},u_1}(0)\right\|_{{\rm hyp},u_1} \leq \kappa_0.$$

To bound the first iteration for v_2, we just have to take into account that it is given by

$$\widetilde{\mathcal{F}}_{{\rm hyp},v_2}(0)(t) = y_2^0 - \int_0^t \left(2\nu_{02} x_2^0 (y_1^0)^2 + \mathcal{O}\left(\left(y_1^0\right)^3 \left(x_2^0\right)^2\right)\right) dt.$$

Then, recalling that $z^0 \in \widehat{\mathcal{N}}_j$,

$$\left|\widetilde{\mathcal{F}}_{{\rm hyp},v_2}(0)(t)\right| \leq 4\nu_{02} x_2^0 (y_1^0)^2 T_j,$$

which gives

$$\left\|\widetilde{\mathcal{F}}_{{\rm hyp},v_2}(0)\right\|_{{\rm hyp},v_2} \leq 4\nu_{02}.$$

Therefore, we can conclude that

$$\left\|\widetilde{\mathcal{F}}(0)\right\|_* \leq \kappa_0$$

for certain constant $\kappa_0 > 0$ independent of δ, σ and j. □

The previous two Propositions show that $\widetilde{\mathcal{F}}_{\mathrm{hyp}}$ is contractive from $B(2\kappa_0) \subset \mathcal{Y}_{\mathrm{hyp}}$ to itself. Therefore, it has a unique fixed point in $B(2\kappa_0) \subset \mathcal{Y}_{\mathrm{hyp}}$ which we denote by w^*. Now it only remains to deduce the bounds for z^f stated in Lemma 5.2. To this end, we use the contractivity of the operator $\widetilde{\mathcal{F}}_{\mathrm{hyp}}$ and we undo the change (5.13). Using the definition of T_j in (5.10), we obtain

$$\begin{aligned} x_2^f =& e^{\sqrt{3}T_j} v_2(T_j) \\ =& \frac{f_2(\sigma)}{x_2^0}\left(x_2^0 + \widetilde{\mathcal{F}}_{\mathrm{hyp},v_2}(w^*)(T_j) - \widetilde{\mathcal{F}}_{\mathrm{hyp},v_2}(0)(T_j)\right) \\ =& f_2(\sigma)\left(1 + \mathcal{O}\left(\left(\sigma \widehat{C}^{(j)}\delta\right)^{1/2} \ln^2(1/\delta)\right)\right) \end{aligned}$$

Analgously, one can see that

$$|y_1^f| \leq K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2}.$$

To obtain the estimates for x_1^f, note that the particular choice that we have done for x_2^* in (5.11) implies that

$$\begin{aligned} |u_1(T_j)| \leq & \left|\widetilde{\mathcal{F}}_{\mathrm{hyp},u_1}(0)(T_j)\right| + \left|\widetilde{\mathcal{F}}_{\mathrm{hyp},u_1}(w^*)(T_j) - \widetilde{\mathcal{F}}_{\mathrm{hyp},u_1}(0)(T_j)\right| \\ \leq & K_\sigma \widehat{C}^{(j)}\delta\left(1 + \mathcal{O}_\sigma\left(\left(\widehat{C}^{(j)}\delta\right)^{1/2} \ln^2(1/\delta)\right)\right). \end{aligned}$$

Then, undoing the change of coordinates (5.13) and using the definition of T_j in (5.10), one obtains

$$|x_1^f| \leq K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2}.$$

Finally, proceeding analogously, and taking into account (5.11) again, one can see that

$$y_2^f = -\frac{f_1(\sigma)}{f_2(\sigma)} \widehat{C}^{(j)}\delta \ln(1/\delta)\left(1 + \mathcal{O}_\sigma\left(\frac{1}{\ln(1/\delta)}\right)\right)$$

which completes the proof of Proposition 5.2. □

Now, it only remains to prove Proposition 5.5.

Proof of Proposition 5.5. To compute the Lipschitz constant we need first upper bounds for $w \in B(2\kappa_0) \subset \mathcal{Y}_{\mathrm{hyp}}$ in the classical supremmum norm $\|\cdot\|_\infty$. They can be deduced from the definition of the norms $\|\cdot\|_{\mathrm{hyp},*}$ in (5.17) and the fact that $z^0 \in \widetilde{\mathcal{N}}^{(j)}$ (see (5.8)). Then, we have that

$$\begin{aligned} |u_1| &\leq K_\sigma \widehat{C}^{(j)}\delta \ln(1/\delta) \\ |v_1| &\leq K\sigma \\ |u_2| &\leq K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2} \\ |v_2| &\leq K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2} \ln(1/\delta). \end{aligned} \tag{5.19}$$

where $K > 0$ is a constant independent of σ.

We use these bounds to obtain the Lipschitz constant. We start by computing the Lipschitz constant of $\widetilde{\mathcal{F}}_{\mathrm{hyp},v_1}=\mathcal{F}_{\mathrm{hyp},v_1}$ and $\widetilde{\mathcal{F}}_{\mathrm{hyp},u_2}=\mathcal{F}_{\mathrm{hyp},u_2}$ and then we will compute the other two.

Using the properties of R_{hyp,y_1} given in Lemma 5.1, (5.12) and the just obtained bounds, one can easily see that

$$
\begin{aligned}
\left|\mathcal{F}_{\mathrm{hyp},v_1}(u,v)-\mathcal{F}_{\mathrm{hyp},v_1}(u',v')\right| \leq & \int_0^{T_j}\mathcal{O}\left(uv\right)\sum_{i=1,2}|v_i-v_i'|dt+\int_0^{T_j}\mathcal{O}\left(v^2\right)\sum_{i=1,2}|u_i-u_i'|dt\\
\leq & K_\sigma\left(\widehat{C}^{(j)}\delta\right)^{1/2}\ln(1/\delta)\sum_{i=1,2}\|v_i-v_i'\|_\infty\\
& +K_\sigma\ln(1/\delta)\sum_{i=1,2}\|u_i-u_i'\|_\infty\\
\leq & K_\sigma\left(\widehat{C}^{(j)}\delta\right)^{1/2}\ln(1/\delta)\sum_{i=1,2}\|v_i-v_i'\|_{\mathrm{hyp},v_i}\\
& +K_\sigma\left(\widehat{C}^{(j)}\delta\right)^{1/2}\ln(1/\delta)\sum_{i=1,2}\|u_i-u_i'\|_{\mathrm{hyp},u_i}.
\end{aligned}
$$

Note that we are abusing notation since inside the $\mathcal{O}(\cdot)$ the dependence of the size on (u,v) means both dependence on (u,v) and (u',v'). We do not write the full dependence since both terms have the same size. Applying the norms defined in (5.17), we get

$$
\left\|\mathcal{F}_{\mathrm{hyp},v_1}(u,v)-\mathcal{F}_{\mathrm{hyp},v_1}(u',v')\right\|_{\mathrm{hyp},v_1}\leq K_\sigma\left(\widehat{C}^{(j)}\delta\right)^{1/2}\ln(1/\delta)\|(u,v)-(u',v')\|_*.
$$

Now we bound the Lipschitz constant of $\mathcal{F}_{\mathrm{hyp},u_2}$. Proceeding as in the previous case one obtains

$$
\begin{aligned}
\left|\mathcal{F}_{\mathrm{hyp},u_2}(u,v)-\mathcal{F}_{\mathrm{hyp},u_2}(u',v')\right| \leq & \int_0^{T_j}\mathcal{O}\left(uv\right)\sum_{i=1,2}|u_i-u_i'|dt+\int_0^{T_j}\mathcal{O}\left(u^2\right)\sum_{i=1,2}|v_i-v_i'|dt\\
\leq & K_\sigma\left(\widehat{C}^{(j)}\delta\right)^{1/2}\ln(1/\delta)\sum_{i=1,2}\|u_i-u_i'\|_\infty\\
& +K_\sigma\widehat{C}^{(j)}\delta\ln(1/\delta)\sum_{i=1,2}\|v_i-v_i'\|_\infty.\\
\leq & K_\sigma\widehat{C}^{(j)}\delta\ln(1/\delta)\sum_{i=1,2}\|u_i-u_i'\|_{\mathrm{hyp},u_i}\\
& +K_\sigma\widehat{C}^{(j)}\delta\ln(1/\delta)\sum_{i=1,2}\|v_i-v_i'\|_{\mathrm{hyp},v_i}
\end{aligned}
$$

and thus

$$
\left\|\mathcal{F}_{\mathrm{hyp},u_2}(u,v)-\mathcal{F}_{\mathrm{hyp},u_2}(u',v')\right\|_{\mathrm{hyp},u_2}\leq K_\sigma\left(\widehat{C}^{(j)}\delta\right)^{1/2}\ln(1/\delta)\|(u,v)-(u',v')\|_*.
$$

To bound the Lipschitz constant of $\widetilde{\mathcal{F}}_{\mathrm{hyp},u_1}$ we use its definition in (5.16). First we study $\mathcal{F}_{\mathrm{hyp},u_1}(w)-\mathcal{F}_{\mathrm{hyp},u_1}(w')$. We proceed as for $\mathcal{F}_{\mathrm{hyp},u_2}$ but we have to be more accurate. We obtain

$$
\begin{aligned}
\left|\mathcal{F}_{\mathrm{hyp},u_1}(u,v)-\mathcal{F}_{\mathrm{hyp},u_1}(u',v')\right| \leq & \int_0^{T_j} \mathcal{O}\left(uv\right) \sum_{i=1,2} |u_i-u_i'|dt + \int_0^{T_j} \mathcal{O}\left(u^2\right) \sum_{i=1,2} |v_i-v_i'|dt \\
\leq & K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2} \ln(1/\delta) \sum_{i=1,2} \|u_i-u_i'\|_\infty \\
& + K_\sigma \widehat{C}^{(j)}\delta \ln(1/\delta) \sum_{i=1,2} \|v_i-v_i'\|_\infty \\
\leq & K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2} \widehat{C}^{(j)}\delta \ln^2(1/\delta) \|u_1-u_1\|_{\mathrm{hyp},u_1} \\
& + K_\sigma \widehat{C}^{(j)}\delta \ln(1/\delta) \|u_2-u_2'\|_{\mathrm{hyp},u_2} \\
& + K_\sigma \widehat{C}^{(j)}\delta \ln(1/\delta) \|v_1-v_1'\|_{\mathrm{hyp},v_1} \\
& + K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2} \widehat{C}^{(j)}\delta \ln^2(1/\delta) \|v_2-v_2'\|_{\mathrm{hyp},v_2}.
\end{aligned}
$$

Thus, taking into account that for δ small enough,

$$
\sup_{t\in[0,T_j(x_2^0)]} \left| \frac{1}{-\widehat{C}^{(j)}\delta\ln(1/\delta)+2\nu_{02}f_1(\sigma)\left(x_2^*\right)^2 t+\widehat{C}^{(j)}\delta} \right| \leq \frac{2}{\widehat{C}^{(j)}\delta},
$$

one can deduce that

$$
\begin{aligned}
\left\|\mathcal{F}_{\mathrm{hyp},u_1}(u,v)-\mathcal{F}_{\mathrm{hyp},u_1}(u',v')\right\|_{\mathrm{hyp},u_1} \leq & K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2} \ln^2(1/\delta) \|u_1-u_1\|_{\mathrm{hyp},u_1} \\
& + K_\sigma \ln(1/\delta) \|u_2-u_2'\|_{\mathrm{hyp},u_2} \\
& + K_\sigma \ln(1/\delta) \|v_1-v_1'\|_{\mathrm{hyp},v_1} \\
& + K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2} \ln^2(1/\delta) \|v_2-v_2'\|_{\mathrm{hyp},v_2}.
\end{aligned}
$$

Therefore, to obtain the Lipschitz constant for $\widetilde{\mathcal{F}}_{\mathrm{hyp},u_1}$, it only remains to use its definition in (5.16) and the Lipschitz constants already obtained for $\mathcal{F}_{\mathrm{hyp},v_1}$ and $\mathcal{F}_{\mathrm{hyp},u_2}$ to obtain

$$
\left\|\widetilde{\mathcal{F}}_{\mathrm{hyp},u_1}(u,v)-\widetilde{\mathcal{F}}_{\mathrm{hyp},u_1}(u',v')\right\|_{\mathrm{hyp},u_1} \leq K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2} \ln^2(1/\delta) \|(u,v)-(u',v')\|_*.
$$

Proceeding analogously, one can see also that

$$
\left\|\widetilde{\mathcal{F}}_{\mathrm{hyp},v_2}(u,v)-\widetilde{\mathcal{F}}_{\mathrm{hyp},v_2}(u',v')\right\|_{\mathrm{hyp},v_2} \leq K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2} \ln(1/\delta) \|(u,v)-(u',v')\|_*.
$$

This completes the proof. □

6 The local map: proof of Lemma 4.7

Analysis of Section 5 describes dynamics of the Hyperbolic Toy Model (5.1). Now we add the elliptic modes and consider the whole vector field (4.14). Our goal is to study the map $\mathcal{B}^j_{\mathrm{loc}}$. The key point of this study is that the elliptic modes remain almost constant through the saddle map and do not make much influence on the hyperbolic ones. In other words, there is *an almost* *product structure.* This allows us to extend the results obtained for the hyperbolic toy model (5.1) in Section 5 to the general system.

As a first step we perform the change obtained in Lemma 5.1 by means of a normal form procedure for the Hyperbolic Toy Model (5.1). The proof of this lemma is straightforward taking into account the form of the vector field (4.14) and the properties of Ψ_{hyp} given in Lemma 5.1.

Lemma 6.1. *Let Ψ_{hyp} be the map defined in Lemma 5.1. Then an application of the change of coordinates*

$$(p_1, q_1, p_2, q_2, c) = (\Psi_{\mathrm{hyp}}(x_1, y_1, x_2, y_2), c), \tag{6.1}$$

to the vector field (4.14) *leads to a vector field of the form*

$$\begin{aligned}
\dot{z} &= Dz + R_{\mathrm{hyp}}(z) + R_{\mathrm{mix},z}(z, c)\\
\dot{c}_k &= ic_k + \mathcal{Z}_{\mathrm{ell},c_k}(c) + R_{\mathrm{mix},c}(z, c),
\end{aligned}$$

where z denotes $z = (x_1, y_1, x_2, y_2)$, $D = \mathrm{diag}(\sqrt{3}, -\sqrt{3}, \sqrt{3}, -\sqrt{3})$, R_{hyp} has been given in Lemma 5.1, $\mathcal{Z}_{\mathrm{ell},c_k}$ is defined in (4.19)*, and $R_{\mathrm{mix},z}$ and R_{mix,c_k} are defined as*

$$\begin{aligned}
R_{\mathrm{mix},x_1} &= A_{x_1}(z)\overline{c_{j-2}}^2 + \overline{A_{x_1}(z)}c_{j-2}^2 + \frac{\sqrt{3}}{2}\sum_{k\in\mathcal{P}}|c_k|^2\Psi_{x_1}(z)\\
R_{\mathrm{mix},y_1} &= A_{y_1}(z)\overline{c_{j-2}}^2 + \overline{A_{y_1}(z)}c_{j-2}^2 + \frac{\sqrt{3}}{2}\sum_{k\in\mathcal{P}}|c_k|^2\Psi_{y_1}(z)\\
R_{\mathrm{mix},x_2} &= A_{x_2}(z)\overline{c_{j+2}}^2 + \overline{A_{x_2}(z)}c_{j+2}^2 + \frac{\sqrt{3}}{2}\sum_{k\in\mathcal{P}}|c_k|^2\Psi_{x_2}(z)\\
R_{\mathrm{mix},y_2} &= A_{y_2}(z)\overline{c_{j+2}}^2 + \overline{A_{y_2}(z)}c_{j+2}^2 + \frac{\sqrt{3}}{2}\sum_{k\in\mathcal{P}}|c_k|^2\Psi_{y_2}(z)\\
R_{\mathrm{mix},c_k} &= i\sqrt{3}c_kP(z) \quad \textit{for } m \neq j \pm 2\\
R_{\mathrm{mix},c_{j\pm 2}} &= i\sqrt{3}c_{j\pm 2}P(z) - i\overline{c}_{j\pm 2}Q_{\pm}(z)
\end{aligned}$$

where $\Psi_{\mathrm{hyp},z}$ are the functions defined in Lemma 5.1, A_z satisfy

$$A_{x_i} = \mathcal{O}(x_i, y_i) \quad \textit{and} \quad A_{y_i} = \mathcal{O}(x_i, y_i)$$

and P and $Q_\pm$ satisfy

$$P(z) = \mathcal{O}\left(x_1y_1, x_2y_2, z_1^2z_2^2\right), \quad Q_-(z) = \mathcal{O}\left(x_1, y_1\right) \quad \textit{and} \quad Q_+(z) = \mathcal{O}\left(x_2, y_2\right).$$

One can easily see that for this system there is a rather strong interaction between the hyperbolic and the elliptic modes due to the terms R_{mix,x_i} and R_{mix,y_i}. The importance of these terms can be seen as follows. The manifold $\{x = 0, y = 0\}$ is normally hyperbolic [Fen74, Fen77, HPS77] for the linear truncation of the vector field obtained in Lemma 6.1 and its stable and unstable manifolds are defined as $\{x = 0\}$ and $\{y = 0\}$. For the full vector field, the manifold $\{x = 0, y = 0\}$ is persistent. Moreover it is still normally hyperbolic thanks to [Fen74, Fen77, HPS77]. Nevertheless, the associated invariant manifolds deviate from $\{x = 0\}$ and $\{y = 0\}$ due to the terms R_{mix,x_i} and R_{mix,y_i}. To overcome this problem, we slightly modify the change (6.1) to straighten these invariant manifolds completely.

Lemma 6.2. *There exist a change of coordinates of the form*

$$(p_1, q_1, p_2, q_2, c) = (\Psi(x_1, y_1, x_2, y_2, c), c) = (x_1, y_1, x_2, y_2, c) + \left(\widetilde{\Psi}(x_1, y_1, x_2, y_2, c), 0\right) \tag{6.2}$$

which transforms the vector field (4.14) *into a vector field of the form*

$$\begin{aligned} \dot{z} &= Dz + R_{\mathrm{hyp}}(z) + \widetilde{R}_{\mathrm{mix},z}(z, c) \\ \dot{c}_k &= ic_k + \mathcal{Z}_{\mathrm{ell},c_k}(c) + \widetilde{R}_{\mathrm{mix},c_k}(z, c), \end{aligned} \tag{6.3}$$

where R_{hyp} and $\mathcal{Z}_{\mathrm{ell}}$ are the functions defined in (5.3) *and* (4.19) *respectively, and*

$$\begin{aligned}
\widetilde{R}_{\mathrm{mix},x_1} &= B_{x_1}(z,c)\overline{c_{j-2}}^2 + \overline{B_{x_1}(z,c)}c_{j-2}^2 + \frac{\sqrt{3}}{2}\sum_{k\in\mathcal{P}}|c_k|^2 C_{x_1}(z,c) \\
\widetilde{R}_{\mathrm{mix},y_1} &= B_{y_1}(z,c)\overline{c_{j-2}}^2 + \overline{B_{y_1}(z,c)}c_{j-2}^2 + \frac{\sqrt{3}}{2}\sum_{k\in\mathcal{P}}|c_k|^2 C_{y_1}(z,c) \\
\widetilde{R}_{\mathrm{mix},x_2} &= B_{x_2}(z,c)\overline{c_{j+2}}^2 + \overline{B_{x_2}(z,c)}c_{j+2}^2 + \frac{\sqrt{3}}{2}\sum_{k\in\mathcal{P}}|c_k|^2 C_{x_2}(z,c) \\
\widetilde{R}_{\mathrm{mix},y_2} &= B_{y_2}(z,c)\overline{c_{j+2}}^2 + \overline{B_{y_2}(z,c)}c_{j+2}^2 + \frac{\sqrt{3}}{2}\sum_{k\in\mathcal{P}}|c_k|^2 C_{y_2}(z,c) \\
\widetilde{R}_{\mathrm{mix},c_k} &= i\sqrt{3}c_k\widetilde{P}(z,c) \quad \text{for } k \neq j \pm 2 \\
\widetilde{R}_{\mathrm{mix},c_{j\pm2}} &= i\sqrt{3}c_{j\pm2}\widetilde{P}(z,c) - i\overline{c}_{j\pm2}\widetilde{Q}_\pm(z,c),
\end{aligned}$$

where the functions B_z and C_z satisfy

$$\begin{aligned}
B_{x_1}(z,c) &= \mathcal{O}(x_1 + y_1x_2z_2) & \qquad B_{x_2}(z,c) &= \mathcal{O}(x_2 + y_2x_1z_1) \\
B_{y_1}(z,c) &= \mathcal{O}(y_1 + x_1y_2z_2) & B_{y_2}(z,c) &= \mathcal{O}(y_2 + x_2y_1z_1) \\
C_{x_1}(z,c) &= \mathcal{O}(x_1 + y_1x_2z_2) & C_{x_2}(z,c) &= \mathcal{O}(x_2 + y_2x_1z_1) \\
C_{y_1}(z,c) &= \mathcal{O}(y_1 + x_1y_2z_2) & C_{y_2}(z,c) &= \mathcal{O}(y_2 + x_2y_1z_1)
\end{aligned}$$

and $\widetilde{P}$ and $\widetilde{Q}_\pm$ satisfy

$$\widetilde{P}(z,c) = \mathcal{O}\left(x_1y_1, x_2y_2, z_1^2z_2^2\right), \quad \widetilde{Q}_-(z,c) = \mathcal{O}(x_1, y_1) \quad \text{and} \quad \widetilde{Q}_+(z) = \mathcal{O}(x_2, y_2).$$

Moreover, the function $\widetilde{\Psi}$ satisfies

$$\begin{aligned}
\widetilde{\Psi}_{x_1} &= \mathcal{O}\left(x_1^3, x_1y_1, x_1(x_2^2 + y_2^2), y_1y_2(x_2 + y_2), c_{j-2}^2y_1, \sum_{k\in\mathcal{P}}|c_k|^2y_1y_2^2\right) \\
\widetilde{\Psi}_{y_1} &= \mathcal{O}\left(y_1^3, x_1y_1, y_1(x_2^2 + y_2^2), x_1x_2(x_2 + y_2), c_{j-2}^2x_1, \sum_{k\in\mathcal{P}}|c_k|^2x_1x_2^2\right) \\
\widetilde{\Psi}_{x_2} &= \mathcal{O}\left(x_2^3, x_2y_2, x_2(x_1^2 + y_1^2), y_1y_2(x_1 + y_1), c_{j+2}^2y_1, \sum_{k\in\mathcal{P}}|c_k|^2y_2y_1^2\right) \\
\widetilde{\Psi}_{y_2} &= \mathcal{O}\left(y_2^3, x_2y_2, y_2(x_1^2 + y_1^2), x_1x_2(x_1 + y_1), c_{j+2}^2x_1, \sum_{k\in\mathcal{P}}|c_k|^2x_2x_1^2\right).
\end{aligned}$$

Proof. It is enough to compose two change of coordinates. The first change is the change (6.2) considered in Lemma 6.1. The second one is the one which straightens the invariant manifolds of a normally hyperbolic invariant manifold [Fen74, Fen77, HPS77]. Then, to obtain the required estimates, it suffices to combine Lemmas 5.1 and 6.1 with the standard results about normally hyperbolic invariant manifolds. □

After performing this change of coordinates, the stable and unstable invariant manifolds of $\{x = 0, y = 0\}$ are straightened. This will facilitate the study of the transition map close to the saddle.

As we have done in Section 5, we define a set $\widehat{\mathcal{V}}_j$ such that

$$\Upsilon(\mathcal{V}_j) \subset \widehat{\mathcal{V}}_j, \tag{6.4}$$

where $\mathcal{V}_j$ is the set defined in Lemma 4.7 and Υ is the inverse of the coordinate change Ψ given in Lemma 6.2. Then, we will apply the flow $\widehat{\Phi}^t$ associated to the vector field (6.3) to points in $\widehat{\mathcal{V}}_j$. To obtain the inclusion (6.4) we define the function $g_{\mathcal{I}_j}(p_2, q_2, \sigma, \delta)$ involved in the definition of $\mathcal{V}_j$.

Define the set

$$\widehat{\mathcal{V}}_j = \mathbb{D}_1^1 \times \ldots \times \mathbb{D}_j^{j-2} \times \widehat{\mathcal{N}}_j \times \mathbb{D}_j^{j+2} \times \ldots \times \mathbb{D}_j^N,$$

where $\widehat{\mathcal{N}}_j$ is the set defined in (5.8) and $\mathbb{D}_j^k$ are defined as

$$\begin{aligned} \mathbb{D}_j^k &= \left\{ |c_k| \leq M_{\text{ell},\pm} \delta^{(1-r)/2} \right\} \quad \text{for } k \in \mathcal{P}_j^{\pm} \\ \mathbb{D}_j^{j\pm 2} &= \left\{ |c_{j\pm 2}| \leq M_{\text{adj},\pm} \left(\widehat{C}^{(j)} \delta \right)^{1/2} \right\}. \end{aligned}$$

Define the function $g_{\mathcal{I}_j}(p_2, q_2, \sigma, \delta)$ involved in the definition of the set $\mathcal{V}_j$ as

$$g_{\mathcal{I}_j}(p_2, q_2, \sigma, \delta) = p_2 + a_p(\sigma) p_2 + a_q(\sigma) q_2 - x_2^* \tag{6.5}$$

where x_2^* is the constant defined in (5.11) and

$$\begin{aligned} a_p(\sigma) &= \partial_{p_2} \widetilde{\Upsilon}_{p_2}(0, \sigma, 0, 0, 0) \\ a_q(\sigma) &= \partial_{q_2} \widetilde{\Upsilon}_{p_2}(0, \sigma, 0, 0, 0), \end{aligned}$$

where $\Upsilon = \text{Id} + \widetilde{\Upsilon}$ is the inverse of the change Ψ given in Lemma 6.2.

Lemma 6.3. *With the above notations for δ small enough condition* (6.4) *is satisfied.*

Proof. It is a straightforward consequence of Lemmas 5.1 and 6.2. □

After straightening the invariant manifold, next lemma studies the saddle map in the transformed variables for points belonging to $\mathcal{V}_j$.

Lemma 6.4. *Let us consider the flow $\widehat{\Phi}_t$ associated to* (6.3) *and a point $(z^0, c^0) \in \widehat{\mathcal{V}}_j$. Then for δ and σ small enough, the point*

$$\left(z^f, c^f\right) = \widehat{\Phi}_{T_j}\left(z^0, c^0\right),$$

where $T_j = T_j(x_2^0)$ *is the time defined in* (5.10)*, satisfies*

$$\begin{aligned}
|x_1^f| &\leq K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2}\\
|y_1^f| &\leq K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2}\\
|x_2^f - f_2(\sigma)| &\leq K_\sigma \delta^{r'}\\
\left|y_2^f + \frac{f_1(\sigma)}{f_2(\sigma)}\widehat{C}^{(j)}\delta\ln(1/\delta)\right| &\leq \frac{f_1(\sigma)}{f_2(\sigma)}\delta.
\end{aligned}$$

and

$$\begin{aligned}
\left|c_k^f - c_k^0 e^{iT_j}\right| &\leq K_\sigma \delta^{(1-r)/2+r'} \quad \text{for } k \in \mathcal{P}_j^\pm\\
\left|c_{j\pm2}^f - c_{j\pm2}^0 e^{iT_j}\right| &\leq 2M_{\mathrm{adj},\pm}\sigma\left(\widehat{C}^{(j)}\delta\right)^{1/2}.
\end{aligned}$$

We postpone the proof of this lemma to Section 6.1.

Now, to complete the proof of Lemma 4.7 we need two steps.

The first is to undo the change of coordinates performed in Lemma 6.2 to express the estimates of the saddle map in the original variables.

The second step is to adjust the time so that the image belongs to the section Σ_j^{out}. These two final steps are done in the next two following lemmas.

Concerning the first step, recall that the change of variables Ψ defined in Lemma 6.2 does not change the elliptic variables, and therefore it only affects the hyperbolic ones.

Lemma 6.5. *Let us consider the flow* Φ_t *associated to* (4.14) *and a point* $(p^0, q^0, c^0) \in \widehat{\mathcal{V}}_j$*. Then for* δ *and* σ *small enough, the point*

$$\left(p^f, q^f, c^f\right) = \Phi_{T_j}\left(p^0, q^0, c^0\right),$$

where T_j *is the time defined in* (5.10)*, satisfies*

$$\begin{aligned}
|p_1^f| &\leq K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2}\\
|q_1^f| &\leq K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2}\\
|p_2^f - \sigma| &\leq K_\sigma \delta^{r'}\\
|q_2^f + \widetilde{C}^{(j)}\delta\ln(1/\delta)| &\leq \widetilde{C}^{(j)}\ \delta\ K_\sigma.
\end{aligned}$$

for certain constant $\widetilde{C}^{(j)}$ *satisfying* $C^{(j)}/2 \leq \widetilde{C}^{(j)} \leq 2C^{(j)}$ *and*

$$\begin{aligned}
\left|c_k^f - c_k^0 e^{iT_j}\right| &\leq K_\sigma \delta^{(1-r)/2+r'} \quad \text{for } m \in \mathcal{P}^\pm\\
\left|c_{j\pm2}^f - c_{j\pm2}^0 e^{iT_j}\right| &\leq 2M_{\mathrm{adj},\pm}\sigma\left(\widehat{C}^{(j)}\delta\right)^{1/2}.
\end{aligned}$$

Proof. In Lemma 6.2 we have defined the change Ψ which relates the two sets of coordinates by

$$\left(p_1^f, q_1^f, p_2^f, q_2^f, c^f\right) = \left(\Psi\left(x_1^f, y_1^f, x_2^f, y_2^f, c^f\right), c^f\right).$$

Then, taking into account the properties of the change Ψ stated in this lemma, one can easily see that from the estimates obtained in Lemma 6.4, one can deduce the estimates stated in Lemma 6.5. First recall that the change Ψ does not modify the elliptic modes and therefore we only need to deal with the hyperbolic ones.

Using the properties of Ψ and modifying slightly K_σ, it is easy to see that for δ small enough,

$$\begin{aligned}|p_1^f| \leq & K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2}\\ |q_1^f| \leq & K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2}.\end{aligned}$$

To obtain the estimates for p_2 it is enough to recall the definition of $f_2(\sigma)$ in (5.9). For the estimates for q_2, it is enough to see that from the properties of Ψ and the estimates for z^f one can deduce that

$$q_2 = \partial_{x_2}\Psi_{x_2}(0,0,\sigma,0)x_2 + \mathcal{O}_\sigma\left(\widehat{C}^{(j)}\delta\right).$$

Therefore, we can define a constant $\widetilde{C}^j$ such that the estimate for q_2 is satisfied. □

Once we have obtained good estimates for the approximate time map in the original variables, we adjust it to obtain image points belonging to the section Σ_j^{out}.

Lemma 6.6. *Let us consider a point $\left(p^f, q^f, c^f\right) \in \Phi^{T_j}(\mathcal{V}_j)$, where Φ^t is the flow of (4.14), T_j is the time defined in (5.10) and $\mathcal{V}_j$ is the set considered in Theorem 5.*

Then, there exists a time T', which depends on the point $\left(p^f, q^f, c^f\right)$, such that

$$(p^*, q^*, c^*) = \Phi^{T'}\left(p^f, q^f, c^f\right) \in \Sigma_j^{\mathrm{out}}.$$

Moreover, there exists a constant K_σ such that

$$|T'| \leq K_\sigma \delta^r \tag{6.6}$$

and

$$\begin{aligned}\left|c_k^* - c_k^f\right| &\leq K_\sigma \delta^{1-r} \quad \text{for } m \in \mathcal{P}\\ \left|p_1^* - p_1^f\right| &\leq K_\sigma \left(C^{(j)}\delta\right)^{1/2}\delta^{1-r}\\ \left|q_1^* - q_1^f\right| &\leq K_\sigma \left(C^{(j)}\delta\right)^{1/2}\delta^{1-r}\\ p_2 &= \sigma\\ \left|q_2^* - q_2^f\right| &\leq K_\sigma C^{(j)}\delta^{2-r}\ln(1/\delta).\end{aligned}$$

Proof. The proof of this Lemma follows the same lines as the proof of Proposition 7.3. Namely, first we obtain a priori bounds for each variable, which then allow us to obtain more refined estimates. □

To finish the proof of Lemma 4.7, we define $\mathcal{U}_j = \mathcal{B}^j_{\mathrm{loc}}(\mathcal{V}_j)$ and we check that this set has a $\widetilde{\mathcal{I}}_j$-product-like structure for a multiindex $\widetilde{\mathcal{I}}_j$ satisfying the properties stated in Lemma 4.7 (see Definition 4.6). Indeed, from the results obtained in Lemmas 6.5 and 6.6 and recalling that by the hypotheses of Lemma 4.7 we have that $M_{\mathrm{hyp}}^{(j)} \geq 1$, it is easy to see that one can define a constant K_σ so that if we consider the constants $\widetilde{M}^{(j)}_{\mathrm{ell},\pm}$, $\widetilde{M}^{(j)}_{\mathrm{adj},\pm}$ and $\widetilde{M}^{(j)}_{\mathrm{hyp}}$ defined in Lemma 4.7 and the constant $\widetilde{C}^{(j)}$ given in Lemma 6.5, the set $\mathcal{U}_j = \mathcal{B}^j_{\mathrm{loc}}(\mathcal{V}_j)$ satisfies condition **C1** stated in Definition 4.6.

Thus, it only remains to check that the set $\mathcal{U}_j$ also satisfies condition **C2** of Definition 4.6. First we check the part of the condition **C2** concerning the elliptic modes. Indeed, from the estimates for the non-neighbor and adjacent elliptic modes given in Lemma 6.5 and 6.6, one can easily see that for any fixed values for the hyperbolic modes, if one takes the constants $\widetilde{m}^{(j)}_{\mathrm{ell}}$, $\widetilde{m}^{(j)}_{\mathrm{adj}}$ given in Lemma 4.7, the image of the elliptic modes contains disks as stated in Definition 4.6. Then, it only remains to check that the inclusion condition is also satisfied for the variable q_2. From the proof of Lemma 6.4 given in Section 6.1, one can easily deduce that the image in the y_2 variable contains an interval of length $\mathcal{O}(\widehat{C}^{(j)}\delta)$ and whose points are of size smaller than $2\widehat{C}^{(j)}\delta\ln(1/\delta)$. Then, when we undo the normal form change of coordinates (Lemma 6.5), this interval is only modified slightly but keeping still a length of order $\mathcal{O}(\widehat{C}^{(j)}\delta)$. Thus taking into account the constant $\widetilde{C}^{(j)}$ given Lemma 6.5 and the results of Lemma 6.6, we can obtain a constant $\widetilde{m}^{(j)}_{\mathrm{hyp}}$ so that condition **C2** is satisfied.

Finally, it only remains to obtain upper bounds for the time spent by the map $\mathcal{B}^j_{\mathrm{loc}}$. To this end it is enough to recall that the time spent is the sum of the time T_j defined in (5.10), which has been bounded in (5.12), and the time T' given in Lemma 6.6, which has been bounded in (6.6). Thus, taking into accounts these two bounds we obtain the bound for the time spent by $\mathcal{B}^j_{\mathrm{loc}}$ given in Lemma 4.7. This finishes the proof of Lemma 4.7.

6.1 Proof of Lemma 6.4

As we have done in the Section 5, we make variation of constants to set up a fixed point argument. Namely, we consider

$$x_i = e^{\sqrt{3}t}u_i,\ y_i = e^{-\sqrt{3}t}v_i,\ c_k = e^{it}d_k$$

and then we obtain the integral equation

$$\begin{aligned}
u_i &= x_i^0 + \int_0^{T_j} e^{-\sqrt{3}t}\left(R_{\mathrm{hyp},x_i}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}\right) + \widetilde{R}_{\mathrm{mix},x_i}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}, de^{it}\right)\right)dt\\
v_i &= y_i^0 + \int_0^{T_j} e^{\sqrt{3}t}\left(R_{\mathrm{hyp},y_i}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}\right) + \widetilde{R}_{\mathrm{mix},y_i}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}, de^{it}\right)\right)dt\\
d_k &= c_k^0 + \int_0^{T_j} e^{-it}\left(\mathcal{Z}_{\mathrm{ell},c_k}\left(de^{it}\right) + \widetilde{R}_{\mathrm{mix},c_k}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}, de^{it}\right)\right)dt.
\end{aligned} \tag{6.7}$$

Note that the terms $R_{\mathrm{hyp},z}$ are the ones considered in Section 5, and, therefore, we will use the properties of these functions obtained in that section. We use the same integration time T_j in (5.10).

As before, we use (6.7) to set up a fixed point argument in two steps. First we define

$\mathcal{G} = (\mathcal{G}_{\mathrm{hyp}}, \mathcal{G}_{\mathrm{ell}})$ as

$$\begin{aligned}
\mathcal{G}_{\mathrm{hyp},u_i}(u,v,d) &= x_i^0 + \int_0^{T_j} e^{-\sqrt{3}t} \left(R_{\mathrm{hyp},x_i}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}\right) + \widetilde{R}_{\mathrm{mix},x_i}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}, de^{it}\right)\right) dt \\
&= \mathcal{F}_{\mathrm{hyp},u_i}(u,v) + \int_0^{T_j} e^{-\sqrt{3}t} \widetilde{R}_{\mathrm{mix},x_i}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}, de^{it}\right) dt \\
\mathcal{G}_{\mathrm{hyp},v_i}(u,v,d) &= y_i^0 - \int_0^{T_j} e^{\sqrt{3}t} \left(R_{\mathrm{hyp},y_i}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}\right) + \widetilde{R}_{\mathrm{mix},x_i}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}, de^{it}\right)\right) dt \\
&= \mathcal{F}_{\mathrm{hyp},v_i}(u,v) + \int_0^{T_j} e^{\sqrt{3}t} \widetilde{R}_{\mathrm{mix},x_i}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}, de^{it}\right) dt,
\end{aligned}$$

where $\mathcal{F}_{\mathrm{hyp}}$ is the operator defined in (5.15), and

$$\mathcal{G}_{\mathrm{ell},c_k}(u,v,d) = c_k^0 + \int_0^{T_j} e^{-it} \left(\mathcal{Z}_{\mathrm{ell},c_k}\left(de^{it}\right) + \widetilde{R}_{\mathrm{mix},c_k}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}, de^{it}\right)\right) dt.$$

We modify this operator slightly as we have done for $\mathcal{F}_{\mathrm{hyp}}$ in Section 5 to make it contractive. We define

$$\begin{aligned}
\widetilde{\mathcal{G}}_{\mathrm{hyp},u_1}(u,v,d) &= \mathcal{G}_{\mathrm{hyp},u_1}(u_1, \mathcal{G}_{\mathrm{hyp},v_1}(u,v,d), \mathcal{G}_{\mathrm{hyp},u_2}(u,v,d), v_2, d) \\
\widetilde{\mathcal{G}}_{\mathrm{hyp},v_2}(u,v,d) &= \mathcal{G}_{\mathrm{hyp},v_2}(u_1, \mathcal{G}_{\mathrm{hyp},v_1}(u,v,d), \mathcal{G}_{\mathrm{hyp},u_2}(u,v,d), v_2, d).
\end{aligned}$$

We denote the new operator by

$$\widetilde{\mathcal{G}} = \left(\widetilde{\mathcal{G}}_{\mathrm{hyp},u_1}, \mathcal{G}_{\mathrm{hyp},u_2}, \mathcal{G}_{\mathrm{hyp},v_1}, \widetilde{\mathcal{G}}_{\mathrm{hyp},v_2}, \mathcal{G}_{\mathrm{ell}} \right), \tag{6.8}$$

whose fixed points coincide with those of $\mathcal{G}$.

We extend the norm defined in (5.17) to incorporate the elliptic modes. To this end, we define

$$\begin{aligned}
\|h\|_{\mathrm{ell},\pm} &= \left(M_{\mathrm{ell},\pm} \delta^{(1-r)/2} \right)^{-1} \|h\|_\infty \\
\|h\|_{\mathrm{adj},\pm} &= M_{\mathrm{adj},\pm}^{-1} \left(\widehat{C}^{(j)} \delta \right)^{-1/2} \|h\|_\infty
\end{aligned}$$

and

$$\|(u,v,d)\|_* = \sup_{\substack{k \in \mathcal{P}_j^\pm \\ i=1,2}} \left\{ \|u_i\|_{\mathrm{hyp},u_i}, \|v_i\|_{\mathrm{hyp},v_i}, \|d_k\|_{\mathrm{ell},\pm}, \|d_{j\pm 2}\|_{\mathrm{adj},\pm} \right\}$$

which, abusing notation, is denoted as the norm in (5.18). We also define the Banach space

$$\mathcal{Y} = \left\{ (u,v,d) : [0,T] \to \mathbb{C}^{N-3} \times \mathbb{R}^4 ; \|(u,v,d)\|_* < \infty \right\}.$$

Proceeding as in Section 5, we state the two following propositions, from which one can easily deduce the contractivity of $\widetilde{\mathcal{G}}$. The proof of the first one is straightforward taking into account the definition of $\widetilde{\mathcal{G}}$ and Lemma 5.4 and the proof of the second one is deferred to end of the section.

Proposition 6.7. *Let us consider the operator* $\widetilde{\mathcal{G}}$ *defined in* (6.8)*. Then, the components of* $\widetilde{\mathcal{G}}(0)$ *are given by*

$$\begin{aligned}
\widetilde{\mathcal{G}}_{\mathrm{hyp},u_1}(0) &= \widetilde{\mathcal{F}}_{\mathrm{hyp},u_i}(0)\\
\widetilde{\mathcal{G}}_{\mathrm{hyp},v_1}(0) &= y_1^0\\
\widetilde{\mathcal{G}}_{\mathrm{hyp},u_2}(0) &= x_2^0\\
\widetilde{\mathcal{G}}_{\mathrm{hyp},v_2}(0) &= \widetilde{\mathcal{F}}_{\mathrm{hyp},v_2}(0)\\
\widetilde{\mathcal{G}}_{\mathrm{ell},c_k}(0) &= c_k^0.
\end{aligned}$$

Thus, there exists a constant $\kappa_1 > 0$ *independent of* σ*,* δ *and* j *such that the operator* $\widetilde{\mathcal{G}}$ *satisfies*

$$\left\|\widetilde{\mathcal{G}}(0)\right\|_* \le \kappa_1.$$

Proposition 6.8. *Let us consider* $w_1, w_2 \in B(2\kappa_1) \subset \mathcal{Y}$*, a constant* r' *satisfying* $0 < r' < 1/2 - 2r$ *and* δ *as defined in Theorem 3. Then taking* σ *small enough and* N *big enough such that* $0 < \delta = e^{-\gamma N} \ll 1$*, there exist a constant* $K_\sigma > 0$ *which is independent of* j *and* N*, but might depend on* σ*, and a constant* K *independent of* j*,* N *and* σ*, such that the operator* $\widetilde{\mathcal{G}}$ *satisfies*

$$\begin{aligned}
&\left\|\widetilde{\mathcal{G}}_{\mathrm{hyp},u_i}(u,v,d) - \widetilde{\mathcal{G}}_{\mathrm{hyp},u_i}(u',v',d')\right\|_{\mathrm{hyp},u_i,v_i} \le\\
&\qquad\qquad \le K_\sigma \delta^{r'} \left\|(u,v,d) - (u',v',d')\right\|_*\\
&\left\|\widetilde{\mathcal{G}}_{\mathrm{hyp},v_i}(u,v,d) - \widetilde{\mathcal{G}}_{\mathrm{hyp},v_i}(u',v',d')\right\|_{\mathrm{hyp},u_i,v_i} \le\\
&\qquad\qquad \le K_\sigma \delta^{r'} \left\|(u,v,d) - (u',v',d')\right\|_*\\
&\left\|\widetilde{\mathcal{G}}_{\mathrm{ell},c_k}(u,v,d) - \widetilde{\mathcal{G}}_{\mathrm{ell},c_k}(u',v',d')\right\|_{\mathrm{ell},\pm} \le\\
&\qquad\qquad \le K_\sigma \delta^{r'} \left\|(u,v,d) - (u',v',d')\right\|_*, \quad \textit{for } m \in \mathcal{P}^\pm\\
&\left\|\widetilde{\mathcal{G}}_{\mathrm{adj},\pm}(u,v,d) - \widetilde{\mathcal{G}}_{\mathrm{adj},\pm}(u',v',d')\right\|_{\mathrm{adj},\pm} \le\\
&\qquad\qquad \le K\sigma \left\|(u,v,d) - (u',v',d')\right\|_*.
\end{aligned}$$

Thus, since $0 < \delta \ll \sigma$*,*

$$\left\|\widetilde{\mathcal{G}}(w_2) - \widetilde{\mathcal{G}}(w_1)\right\|_* \le 2K\sigma \|w_2 - w_1\|_*$$

and therefore, for σ *small enough, it is contractive.*

The previous two propositions show that the operator $\widetilde{\mathcal{G}}$ is contractive. Let us denote by (u^*, v^*, d^*) its unique fixed point in the ball $B(2\kappa_1) \subset \mathcal{Y}$. Now, it only remains to obtain the estimates stated in Lemma 6.4. The estimates for the hyperbolic variables are obtained as in the proof of Lemma 5.2. For the elliptic ones it is enough to take into account that

$$\begin{aligned}
c_k^f = c_k(T_j) &= d_k(T_j) e^{iT_j}\\
&= \mathcal{G}_{\mathrm{ell},c_k}(0)(T_j) e^{iT_j} + \left(\mathcal{G}_{\mathrm{ell},c_k}(u^*,v^*,d^*)(T_j) - \mathcal{G}_{\mathrm{ell},c_k}(0)(T_j)\right) e^{iT_j}\\
&= c_k^0 e^{iT_j} + \left(\mathcal{G}_{\mathrm{ell},c_k}(u^*,v^*,d^*)(T_j) - \mathcal{G}_{\mathrm{ell},c_k}(0)(T_j)\right) e^{iT_j}
\end{aligned}$$

and bound the second term using the Lipschitz constant obtained in Proposition 6.8.

We finish the section by proving Proposition 6.8, which completes the proof of Lemma 6.4.

Proof of Proposition 6.8. As we have done in the proof of Proposition 5.5, first, we stablish bounds for any $(u, v, d) \in B(2\kappa_1) \subset \mathcal{Y}$ in the supremmum norm, which will be used to bound the Lipschitz constant of each component of $\widetilde{\mathcal{G}}$. Indeed, if $(u, v, d) \in B(2\kappa_1) \subset \mathcal{Y}$, it satisfies (5.19) and

$$
\begin{aligned}
|d_k| &\leq K_\sigma \delta^{(1-r)/2} \quad \text{for} \quad k \in \mathcal{P}_j^{\pm} \\
|d_{j\pm 2}| &\leq K_\sigma \left(\widehat{C}^{(j)}\delta\right)^{1/2} \leq K_\sigma \delta^{(1-r)/2}.
\end{aligned}
$$

We bound the Lipschitz constant for each component of $\widetilde{\mathcal{G}}_{\text{ell}}$. We split each component of the operator between the elliptic, hyperbolic and mixed part. We deal first with the elliptic part. It can be seen that for $k \in \mathcal{P}_j^{\pm}$,

$$
\begin{aligned}
\left|\mathcal{Z}_{\text{ell},c_k}\left(d'e^{it}\right) - \mathcal{Z}_{\text{ell},c_k}\left(de^{it}\right)\right| \leq & K_\sigma \delta^{1-r} N(d_k - d'_k) \\
& + K_\sigma \delta \sum_{\ell \in \mathcal{P}_j \setminus \{k\}} (d_\ell - d'_\ell).
\end{aligned}
$$

Therefore,

$$
\left\|\int_0^{T_j} e^{-it}\left(\mathcal{Z}_{\text{ell},c_k}\left(de^{it}\right) - \mathcal{Z}_{\text{ell},c_k}\left(d'e^{it}\right)\right) dt\right\|_{\text{ell},\pm} \leq K_\sigma \delta^{1-r} N T_j \|(u, v, d) - (u', v', d')\|_*.
$$

Proceeding analogously, one can see also that

$$
\left\|\int_0^{T_j} e^{-it}\left(\mathcal{Z}_{\text{ell},c_{j\pm 2}}\left(de^{it}\right) - \mathcal{Z}_{\text{ell},c_{j\pm 2}}\left(d'e^{it}\right)\right) dt\right\|_{\text{adj},\pm} \leq K_\sigma \delta^{1-r} N T_j \|(u, v, d) - (u', v', d')\|_*.
$$

Now we bound the mixed terms. Proceeding analogously and considering the properties of $\widetilde{R}_{\text{mix},c_k}$ stated in Lemma 6.2, we can see that for $m \neq j \pm 2$,

$$
\begin{aligned}
&\left\|\widetilde{R}_{\text{mix},c_k}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}, de^{it}\right) - \widetilde{R}_{\text{mix},c_k}\left(u'e^{\sqrt{3}t}, v'e^{-\sqrt{3}t}, d'e^{it}\right)\right\|_{\text{ell},\pm} \\
&\quad \leq K_\sigma \widehat{C}^{(j)} \delta \ln^2(1/\delta) \sum_{i=1,2}\left(\|u_i - u'_i\|_{\text{hyp},u_i} + \|v_i - v'_i\|_{\text{hyp},v_i}\right) \\
&\qquad + K_\sigma \widehat{C}^{(j)} \delta \ln^2(1/\delta)\left(\left\|d_k - d'_k\right\|_{\text{ell},\pm} + K_\sigma \delta^{(1-r)/2} \sum_{\ell \in \mathcal{P}_j^{\pm}} \left\|d_\ell - d'_\ell\right\|_{\text{ell},\pm}\right) \\
&\qquad + K_\sigma \widehat{C}^{(j)} \delta^{1+(1-r)/2} \ln^2(1/\delta)\left(\left\|d_{j-2} - d'_{j-2}\right\|_{\text{adj},-} + \left\|d_{j+2} - d'_{j+2}\right\|_{\text{adj},+}\right) \\
&\qquad \leq K_\sigma \widehat{C}^{(j)} \delta \ln^2(1/\delta)\left(1 + K_\sigma N \delta^{(1-r)/2}\right)\left\|(u, v, d) - (u', v', d')\right\|_*.
\end{aligned}
$$

Therefore, using that $\delta = e^{-\gamma N}$ and (5.12),

$$
\begin{aligned}
&\left\|\int_0^{T_j} e^{-it}\left(\widetilde{R}_{\text{mix},c_k}(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}, de^{it}) - \widetilde{R}_{\text{mix},c_k}(u'e^{\sqrt{3}t}, v'e^{-\sqrt{3}t}, d'e^{it})\right) dt\right\|_{\text{ell},\pm} \\
&\qquad \leq K_\sigma \widehat{C}^{(j)} \delta \ln^3(1/\delta) \|(u, v, d) - (u', v', d')\|_*.
\end{aligned}
$$

So, we can conclude that for $m \in \mathcal{P}^{\pm}$,

$$\left\|\mathcal{G}_{\mathrm{ell},c_k}(u,v,d)-\mathcal{G}_{\mathrm{ell},c_k}(u',v',d')\right\|_{\mathrm{ell},\pm} \leq K_\sigma \delta^{1-r}\ln^3(1/\delta)\|(u,v,d)-(u',v',d')\|_*.$$

Proceeding analogously we can bound the Lipschitz constant for $\mathcal{G}_{\mathrm{ell},c_{j\pm 2}}$. We bound it for $m = j-2$, the other case can be done analogously. Here K denotes a generic constant independent of σ. Note that now there is an additional term in $\widetilde{R}_{\mathrm{mix},c_{j-2}}$. This implies that

$$\begin{aligned}
&\left|\widetilde{R}_{\mathrm{mix},c_{j-2}}(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}, de^{it})-\widetilde{R}_{\mathrm{mix},c_{j-2}}(u'e^{\sqrt{3}t}, v'e^{-\sqrt{3}t}, d'e^{it})\right|\\
&\qquad\leq K\sigma M_{\mathrm{adj},-}\left(\widehat{C}^{(j)}\delta\right)^{1/2}e^{-\sqrt{3}t}\sum_{i=1,2}\left(\|u_i-u_i'\|_{\mathrm{hyp},u_i}+\|v_i-v_i'\|_{\mathrm{hyp},v_i}\right)\\
&\qquad+K\sigma M_{\mathrm{adj},-}\left(\widehat{C}^{(j)}\delta\right)^{1/2}e^{-\sqrt{3}t}\left\|d_{j-2}-d_{j-2}'\right\|_{\mathrm{adj},-}\\
&\qquad+K_\sigma M_{\mathrm{adj},-}\left(\widehat{C}^{(j)}\delta\right)^{1/2}\delta^{(1-r)/2}e^{-\sqrt{3}t}\left(\left\|d_{j+2}-d_{j+2}'\right\|_{\mathrm{adj},+}+\sum_{\ell\in\mathcal{P}_j^{\pm}}\left\|d_\ell-d_\ell'\right\|_{\mathrm{ell},\pm}\right)\\
&\qquad\leq K\sigma M_{\mathrm{adj},-}\left(\widehat{C}^{(j)}\delta\right)^{1/2}e^{-\sqrt{3}t}\left\|(u,v,d)-(u',v',d')\right\|_*.
\end{aligned}$$

Therefore, integrating and applying norms, we obtain

$$\begin{aligned}
&\left\|\int_0^{T_j}e^{-it}\left(\widetilde{R}_{\mathrm{mix},c_{j-2}}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}, de^{it}\right)-\widetilde{R}_{\mathrm{mix},c_{j-2}}\left(u'e^{\sqrt{3}t}, v'e^{-\sqrt{3}t}, d'e^{it}\right)\right)dt\right\|_{\mathrm{adj},-}\\
&\qquad\qquad\leq K\sigma\|(u,v,d)-(u',v',d')\|_*,
\end{aligned}$$

which leads to

$$\left\|\mathcal{G}_{\mathrm{ell},c_{j-2}}(u,v,d)-\mathcal{G}_{\mathrm{ell},c_{j-2}}(u',v',d')\right\|_{\mathrm{adj},-}\leq K\sigma\|(u,v,d)-(u',v',d')\|_*.$$

Now we bound the Lipschitz constant for the hyperbolic components of the operator. Note that we only need to bound the terms involving $\widetilde{R}_{\mathrm{mix},z}$ since the other terms of the operator have been bounded in Proposition 5.5. We start with the Lipschitz constants of $\mathcal{G}_{\mathrm{hyp},v_i}$. To this end we bound

$$\begin{aligned}
&\left|\int_0^{T_j}e^{\sqrt{3}t}\left(\widetilde{R}_{\mathrm{mix},y_i}(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}, de^{it})-\widetilde{R}_{\mathrm{mix},y_i}(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}, de^{it})\right)dt\right|\\
&\qquad\leq\int_0^{T_j}\left(\mathcal{O}\left(\sum_{k\in\mathcal{P}_j}|d_k|^2(v_1+v_2)\right)e^{\sqrt{3}t}\left|u_i-u_i'\right|+\mathcal{O}\left(\sum_{k\in\mathcal{P}_j}|d_k|^2\right)\sum|v_i-v_i'|\right)dt\\
&\qquad\quad+\int_0^{T_j}\sum_{k\in\mathcal{P}_j}\mathcal{O}(d_k(v_1+v_2))|d_k-d_k'|dt,
\end{aligned}$$

where we abuse notation concerning the $\mathcal{O}(\cdot)$ as before. Thus, integrating the exponentials and applying norms, one can easily see that

$$\begin{aligned}
&\left|\int_0^{T_j}e^{\sqrt{3}t}\left(\widetilde{R}_{\mathrm{mix},y_i}(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}, de^{it})-\widetilde{R}_{\mathrm{mix},y_i}(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}, de^{it})\right)dt\right|\\
&\qquad\leq K_\sigma N\delta^{1-r}\ln(1/\delta)\|(u,v,d)-(u',v',d')\|_*.
\end{aligned}$$

Therefore, applying norms and using condition on δ from Theorem 3, we obtain

$$\left\|\int_0^{T_j} e^{\sqrt{3}t}\left(\widetilde{R}_{\mathrm{mix},y_i}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}, de^{it}\right) - \widetilde{R}_{\mathrm{mix},y_i}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}, de^{it}\right)\right) dt\right\|_{\mathrm{hyp},v_1}$$
$$\leq K_\sigma \delta^{1-r}\ln^2(1/\delta)\|(u,v,d)-(u'v',d')\|_*$$
$$\left\|\int_0^{T_j} e^{\sqrt{3}t}\left(\widetilde{R}_{\mathrm{mix},y_i}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}, de^{it}\right) - \widetilde{R}_{\mathrm{mix},y_i}\left(ue^{\sqrt{3}t}, ve^{-\sqrt{3}t}, de^{it}\right)\right) dt\right\|_{\mathrm{hyp},v_2}$$
$$\leq K_\sigma \delta^{1/2-2r}\ln(1/\delta)\|(u,v,d)-(u'v',d')\|_*.$$

Then, taking into account the results of Lemma 5.5, one can conclude that

$$\left\|\widetilde{\mathcal{G}}_{\mathrm{hyp},v_1}(u,v,d) - \widetilde{\mathcal{G}}_{\mathrm{hyp},v_1}(u',v',d')\right\|_{\mathrm{hyp},v_1} \leq$$
$$\leq K_\sigma\left(\left(\widehat{C}^{(j)}\delta\right)^{1/2}\ln(1/\delta) + \delta^{1-r}\ln^2(1/\delta)\right)\left\|(u,v,d)-(u',v',d')\right\|_*$$
$$\left\|\widetilde{\mathcal{G}}_{\mathrm{hyp},v_2}(u,v,d) - \widetilde{\mathcal{G}}_{\mathrm{hyp},v_2}(u',v',d')\right\|_{\mathrm{hyp},v_2} \leq$$
$$\leq K_\sigma\left(\left(\widehat{C}^{(j)}\delta\right)^{1/2}\ln(1/\delta) + \delta^{1/2-2r}\ln(1/\delta)\right)\left\|(u,v,d)-(u',v',d')\right\|_*.$$

Proceeding in the same way, one can obtain that

$$\left\|\widetilde{\mathcal{G}}_{\mathrm{hyp},u_1}(u,v,d) - \widetilde{\mathcal{G}}_{\mathrm{hyp},u_1}(u',v',d')\right\|_{\mathrm{hyp},u_1} \leq$$
$$\leq K_\sigma\left(\left(\widehat{C}^{(j)}\delta\right)^{1/2}\ln(1/\delta) + \delta^{1-r}\ln^2(1/\delta)\right)\left\|(u,v,d)-(u',v',d')\right\|_*$$
$$\left\|\widetilde{\mathcal{G}}_{\mathrm{hyp},u_2}(u,v,d) - \widetilde{\mathcal{G}}_{\mathrm{hyp},u_2}(u',v',d')\right\|_{\mathrm{hyp},u_2} \leq$$
$$\leq K_\sigma\left(\left(\widehat{C}^{(j)}\delta\right)^{1/2}\ln(1/\delta) + \delta^{1/2-2r}\ln^2(1/\delta)\right)\left\|(u,v,d)-(u',v',d')\right\|_*.$$

This completes the proof. □

7 The global map: proof of Lemma 4.8

We devote this section to prove Lemma 4.8. The continuous dependence with respect to initial conditions of ordinary differential equations gives for free that the map $\mathcal{B}^j_{\mathrm{glob}}$, defined in (4.36), is well defined for points close enough to the heteroclinic connection defined in (4.3). Nevertheless, to prove Lemma 4.8, we need more accurate estimates.

Recall that the map $\mathcal{B}^j_{\mathrm{glob}}$ is defined in Σ^{out}_j, which is contained in $\mathcal{M}(b) = 1$ (see (4.1)). So, as we have done for $\mathcal{B}^j_{\mathrm{loc}}$, we use the system of coordinates defined in Section 4.1. Recall that the initial section Σ^{out}_j, defined in (4.34), and the final section $\Sigma^{\mathrm{in}}_{j+1}$, defined in (4.26), are expressed in the variables adapted to the j^{th} and $(j+1)^{st}$ saddles respectively. Namely, in the coordinates $(p_1^{(j)}, q_1^{(j)}, p_2^{(j)}, q_2^{(j)}, c^{(j)})$ and $(p_1^{(j+1)}, q_1^{(j+1)}, p_2^{(j+1)}, q_2^{(j+1)}, c^{(j+1)})$ (see Section 7). To simplify the exposition, first we will study the map $\mathcal{B}^j_{\mathrm{glob}}$ expressing both the domain and the image in the variables $(p_1^{(j)}, q_1^{(j)}, p_2^{(j)}, q_2^{(j)}, c^{(j)})$. Then we will express the image of $\mathcal{B}^j_{\mathrm{glob}}$ in the new variables.

To simplify notation we denote the variables adapted to the j^{th} and $(j+1)^{st}$ saddles by

$$(p_1, q_1, p_2, q_2, c) = \left(p_1^{(j)}, q_1^{(j)}, p_2^{(j)}, q_2^{(j)}, c^{(j)}\right)$$

and

$$(\widetilde{p}_1, \widetilde{q}_1, \widetilde{p}_2, \widetilde{q}_2, \widetilde{c}) = \left(p_1^{(j+1)}, q_1^{(j+1)}, p_2^{(j+1)}, q_2^{(j+1)}, c^{(j+1)}\right)$$

and we denote by Θ^j the change of coordinates that relates them, namely

$$(\widetilde{p}_1, \widetilde{q}_1, \widetilde{p}_2, \widetilde{q}_2, \widetilde{c}) = \Theta^j(p_1, q_1, p_2, q_2, c).$$

Lemma 7.1. *The change of coordinates Θ^j is given by*

$$\begin{aligned}
\Theta^j_{\widetilde{c}_k}(p_1, q_1, p_2, q_2, c) &= \frac{\omega p_2 + \omega^2 q_2}{\widetilde{r}} c_k \qquad\qquad \textit{for } k \in \mathcal{P}^{\pm}_{j+1} \cup \{j+3\}\\
\Theta^j_{\widetilde{c}_{j-1}}(p_1, q_1, p_2, q_2, c) &= \frac{\omega p_2 + \omega^2 q_2}{\widetilde{r}} \left(\omega^2 p_1 + \omega q_1\right)\\
\Theta^j_{\widetilde{p}_1}(p_1, q_1, p_2, q_2, c) &= \frac{r}{\widetilde{r}} q_2\\
\Theta^j_{\widetilde{q}_1}(p_1, q_1, p_2, q_2, c) &= \frac{r}{\widetilde{r}} p_2\\
\Theta^j_{\widetilde{p}_2}(p_1, q_1, p_2, q_2, c) &= \operatorname{Re} z + \frac{\sqrt{3}}{3} \operatorname{Im} z\\
\Theta^j_{\widetilde{q}_2}(p_1, q_1, p_2, q_2, c) &= \operatorname{Re} z - \frac{\sqrt{3}}{3} \operatorname{Im} z,
\end{aligned}$$

where $\omega = e^{2\pi i/3}$ and

$$\begin{aligned}
r^2 =& 1 - \sum_{k \neq j-1, j, j+1} |c_k|^2 - (p_1^2 + q_1^2 - p_1 q_1) - (p_2^2 + q_2^2 - p_2 q_2) \qquad (7.1)\\
\widetilde{r}^2 =& p_2^2 + q_2^2 - p_2 q_2\\
z =& \frac{c_{j+2}}{\widetilde{r}} \left(\omega p_2 + \omega^2 q_2\right).
\end{aligned}$$

Proof. We consider a point (p, q, c) and we express it in the new variables. We have to undo the changes (4.8) and (4.5) referred to the saddle j and then apply them again but referred to the saddle $j+1$. The point (p, q, c) has associated variables r (as defined in (7.1)) and θ. We do not need to know the value of θ to deduce the form of the change Θ^j. Indeed, note that if we consider the changes (4.5) and (4.8) for the mode b_{j+1} we have

$$\widetilde{r} e^{i\widetilde{\theta}} = b_{j+1} = c_{j+1} e^{i\theta} = \left(\omega^2 p_2 + \omega q_2\right) e^{i\theta},$$

which implies

$$e^{i(\theta - \widetilde{\theta})} = \frac{\omega p_2 + \omega^2 q_2}{\widetilde{r}}. \qquad (7.2)$$

Using this formula and recalling that $\widetilde{c}_k e^{i\widetilde{\theta}} = b_k = c_k e^{i\theta}$, it is straightforward to deduce the form of $\Theta^j_{\widetilde{c}_k}$ for $k \in \mathcal{P}^{\pm}_{j+1} \cup \{j+3\}$. To deduce the form of $\Theta^j_{\widetilde{p}_1}$ and $\Theta^j_{\widetilde{q}_1}$ it is enough to consider the changes (4.5) and (4.8) for the mode b_j to obtain

$$r e^{i\theta} = b_j = \widetilde{c}_j e^{i\widetilde{\theta}} = \left(\omega^2 \widetilde{p}_1 + \omega \widetilde{q}_1\right) e^{i\widetilde{\theta}}$$

Then, it is enough to use formula (7.2) to obtain $\Theta^j_{\widetilde{p}_1}$ and $\Theta^j_{\widetilde{q}_1}$. The others components can be obtained proceeding in the same way. $\square$

The next step of the proof of Lemma 4.8 is to express the section $\Sigma^{\mathrm{in}}_{j+1}$ in the variables (p_1, q_1, p_2, q_2, c) using the change Θ^j obtained in Lemma 7.1. This is done in the next corollary, which is a straightforward consequence of Lemma 7.1.

Corollary 7.2. *Fix $\sigma > 0$ and define the set*

$$\widetilde{\Sigma}^{\mathrm{in}}_{j+1} = \left(\Theta^j\right)^{-1}\left(\Sigma^{\mathrm{in}}_{j+1} \cap \mathcal{W}_{j+1}\right),$$

where $\Sigma^{\mathrm{in}}_{j+1}$ is the section defined in (4.26) and

$$\mathcal{W}_{j+1} = \left\{|p_1| \leq \eta, |q_1| \leq \eta, |q_2| \leq \eta, |c_k| \leq \eta \text{ for } k \in \mathcal{P}^{\pm}_j \text{ and } k = j \pm 2\right\},$$

Then, for $\eta > 0$ small enough, $\mathcal{W}_{j+1}$ can be expressed as a graph as

$$p_2 = w(p_1, q_1, q_2, c).$$

Moreover, there exist constants κ', κ'' independent of η satisfying

$$0 < \kappa' < \sqrt{1-\sigma^2} < \kappa'' < 1$$

such that, for any $(p_1, q_1, q_2, c) \in \mathcal{W}_{j+1}$, the function w satisfies

$$\kappa' < w(p_1, q_1, q_2, c) < \kappa''.$$

Once we have defined the section $\widetilde{\Sigma}^{\mathrm{in}}_{j+1}$, we can define the map

$$\begin{array}{cccc} \widetilde{\mathcal{B}}^j_{\mathrm{glob}}: & \mathcal{U}_j \subset \Sigma^{\mathrm{out}}_j & \longrightarrow & \widetilde{\Sigma}^{\mathrm{in}}_{j+1} \\ & (p_1, q_1, q_2, c) & \mapsto & \widetilde{\mathcal{B}}^j_{\mathrm{glob}}(p_1, q_1, q_2, c) \end{array}$$

as

$$\widetilde{\mathcal{B}}^j_{\mathrm{glob}} = \Theta_j^{-1} \circ \mathcal{B}^j_{\mathrm{glob}}.$$

We want upper bounds independent of δ and j for the transition time of the correponding orbits for this map. In the variables (p_1, q_1, p_2, q_2, c) the heteroclinic connection (4.3) is simply given by

$$\left(p_1^h(t), q_1^h(t), p_2^h(t), q_2^h(t), c^h(t)\right) = \left(0, 0, \frac{1}{1+e^{2\sqrt{3}(t-t_0)}}, 0, 0\right) \tag{7.3}$$

(see [CKS+10]). Taking t_0 such that

$$\frac{1}{1+e^{2\sqrt{3}t_0}} = \sigma,$$

one can easily see that $p_2^h(2t_0) = \sqrt{1-\sigma^2}$ and $2t_0 \sim \ln(1/\sigma)$. In the new coordinates this point is $(\widetilde{p}_1, \widetilde{q}_1, \widetilde{p}_2, \widetilde{q}_2, \widetilde{c}) = (0, \sigma, 0, 0, 0)$ and thus belongs to the section $\widetilde{q}_1 = \sigma$. Then, thanks to Corollary 7.2, one can easily deduce that the time $T_{\widetilde{\mathcal{B}}^j_{\mathrm{glob}}} = T_{\widetilde{\mathcal{B}}^j_{\mathrm{glob}}}(q_1, p_1, p_2, c)$ spent by the map $\widetilde{\mathcal{B}}^j_{\mathrm{glob}}$ for any point $(q_1, p_1, p_2, c) \in \mathcal{U}_j \subset \Sigma^{\mathrm{out}}_j$ is also independent of δ and j. Recall that the difference between $\widetilde{\mathcal{B}}^j_{\mathrm{glob}}$ and $\mathcal{B}^j_{\mathrm{glob}}$ is just a change of coordinates and therefore the time $T_{\mathcal{B}^j_{\mathrm{glob}}}$ spent by $\mathcal{B}^j_{\mathrm{glob}}$ is the same as $T_{\widetilde{\mathcal{B}}^j_{\mathrm{glob}}}$. Thus, from know on we will only refer to $T_{\mathcal{B}^j_{\mathrm{glob}}}$.

Next step is to study the behavior of the map $\widetilde{\mathcal{B}}^j_{\mathrm{glob}}$. In particular, we want to know the properties of the image set $\widetilde{\mathcal{B}}^j_{\mathrm{glob}}(\mathcal{U}_j)$.

Proposition 7.3. *Let us consider a parameter set $\widetilde{\mathcal{I}}_j$ (as defined in Definition 4.6) and a $\widetilde{\mathcal{I}}_j$-product-like set $\mathcal{U}_j$. Then, there exists a constant $\widetilde{K}_\sigma$ independent of j, N and δ and a constant $D^{(j)}$ satisfying*

$$\widetilde{C}^{(j)}/\widetilde{K}_\sigma \leq D^{(j)} \leq \widetilde{K}_\sigma \widetilde{C}^{(j)},$$

such that the set $\widetilde{\mathcal{B}}^j_{\mathrm{glob}}(\mathcal{U}_j) \subset \widetilde{\Sigma}^{\mathrm{in}}_j$ satisfies the following conditions:

C1

$$\widetilde{\mathcal{B}}^j_{\mathrm{glob}}(\mathcal{U}_j) \subset \widehat{\mathbb{D}}^1_j \times \ldots \times \widehat{\mathbb{D}}^{j-2}_j \times \mathcal{S}_j \times \widehat{\mathbb{D}}^{j+2}_j \times \ldots \times \widehat{\mathbb{D}}^N_j$$

where

$$\widehat{\mathbb{D}}^k_j = \left\{ |c_k| \leq \left(\widetilde{M}^{(j)}_{\mathrm{ell},\pm} + \widetilde{K}_\sigma \delta^{r'} \right) \delta^{(1-r)/2} \right\} \quad \textit{for } k \in \mathcal{P}^\pm_j$$

$$\widehat{\mathbb{D}}^{j\pm 2}_j \subset \left\{ |c_{j\pm 2}| \leq \widetilde{K}_\sigma \widetilde{M}^{(j)}_{\mathrm{adj},\pm} \left(\widetilde{C}^{(j)} \delta \right)^{1/2} \right\},$$

and

$$\begin{aligned}\mathcal{S}_j = \Big\{ &(p_1, q_1, p_2, q_2) \in \mathbb{R}^4 : |p_1|, |q_1| \leq \widetilde{K}_\sigma \widetilde{M}^{(j)}_{\mathrm{hyp}} \left(\widetilde{C}^{(j)} \delta \right)^{1/2}, \\ &p_2 = \sigma, -D^{(j)}\, \delta \left(\ln(1/\delta) - \widetilde{K}_\sigma \right) \leq q_2^{(j)} \leq -D^{(j)}\, \delta \left(\ln(1/\delta) + \widetilde{K}_\sigma \right) \Big\},\end{aligned}$$

C2 *Let us define the projection $\widetilde{\pi}(p, q, c) = (p_2, q_2, c_{j-2}, \ldots, c_N)$. Then,*

$$\left[-D^{(j)}\, \delta \left(\ln(1/\delta) - 1/\widetilde{K}_\sigma \right), -D^{(j)}\, \delta \left(\ln(1/\delta) + 1/\widetilde{K}_\sigma \right) \right] \times \{\sigma\} \times \mathbb{D}^{j+2}_{j,-} \times \ldots \times \mathbb{D}^N_{j,-} \subset \widetilde{\pi} \left(\widetilde{\mathcal{B}}^j_{\mathrm{glob}}(\mathcal{U}_j) \right)$$

where

$$\mathbb{D}^k_{j,-} = \left\{ \left| c^{(j)}_k \right| \leq \left(\widetilde{m}^{(j)}_{\mathrm{ell}} - \widetilde{K}_\sigma \delta^{r'} \right) \delta^{(1-r)/2} \right\} \quad \textit{for } k \in \mathcal{P}^+_j$$

$$\mathbb{D}^{j+2}_{j,-} = \left\{ \left| c^{(j)}_{j+2} \right| \leq \widetilde{m}^{(j)}_{\mathrm{adj}} \left(C^{(j)} \delta \right)^{1/2} / \widetilde{K}_\sigma \right\}.$$

The proof of this proposition is postponed to Section 7.1.

Once we know the properties of the set $\widetilde{\mathcal{B}}^j_{\mathrm{glob}}(\mathcal{U}_j)$, there only remain two final steps. First to deduce analogous properties for the set $\mathcal{B}^j_{\mathrm{glob}}(\mathcal{U}_j) \subset \Sigma^{\mathrm{in}}_{j+1}$. Second, to obtain a parameter set $\mathcal{I}_{j+1}$ and $\mathcal{I}_{j+1}$-product like set $\mathcal{V}_j \subset \Sigma^{\mathrm{in}}_{j+1}$ which satisfies condition (4.40). These two last steps are summarized in the next lemma. Lemma 4.8 follows easily from it.

Lemma 7.4. *Let us consider a parameter set $\mathcal{I}_{j+1}$ whose constants satisfy*

$$D^{(j)}/2 \leq C^{(j+1)} \leq 2D^{(j)}$$

$$0 < m^{(j+1)}_{\mathrm{hyp}} \leq \widetilde{m}^{(j)}_{\mathrm{hyp}}$$

and

$$\begin{aligned}
M_{\mathrm{ell},-}^{(j+1)} &= \max\left\{\widetilde{M}_{\mathrm{ell},-}^{(j)} + \widetilde{K}_\sigma \delta^{r'}, \widetilde{K}_\sigma \widetilde{M}_{\mathrm{adj},-}^{(j)}\right\} \\
M_{\mathrm{ell},+}^{(j+1)} &= \widetilde{M}_{\mathrm{ell},+}^{(j)} + \widetilde{K}_\sigma \delta^{r'} \\
m_{\mathrm{ell}}^{(j+1)} &= \widetilde{m}_{\mathrm{ell}}^{(j)} - \widetilde{K}_\sigma \delta^{r'} \\
M_{\mathrm{adj},+}^{(j+1)} &= \widetilde{m}_{\mathrm{ell},+}^{(j)} + \widetilde{K}_\sigma \delta^{r'} \\
M_{\mathrm{adj},-}^{(j+1)} &= \widetilde{K}_\sigma \widetilde{M}_{\mathrm{hyp}}^{(j)} \\
m_{\mathrm{adj}}^{(j+1)} &= \widetilde{m}_{\mathrm{ell}}^{(j)} + \widetilde{K}_\sigma \delta^{r'} \\
M_{\mathrm{hyp}}^{(j+1)} &= \max\left\{\widetilde{K}_\sigma \widetilde{M}_{\mathrm{adj},+}^{(j)}, \widetilde{K}_\sigma\right\}.
\end{aligned}$$

Then, the set

$$\mathcal{V}_{j+1} = \mathcal{B}_{\mathrm{glob}}^{j}(\mathcal{U}_j) \cap \left\{g_{\mathcal{I}_{j+1}}(p_2, q_2, \sigma, \delta) = 0\right\},$$

where $g_{\mathcal{I}_{j+1}}$ *is the function defined in* (6.5)*, is a* $\mathcal{I}_{j+1}$*-product-like set and satisfies condition* (4.40)

Proof. It is enough to apply the change of coordinates Θ^j given in Lemma 7.1. □

7.1 Proof of Proposition 7.3

We split the proof of Proposition 7.3 in several lemmas, which will give the needed estimates for the different modes. First, let us obtain rough bounds for all the variables, which will be used in the proofs of the forthcoming lemmas. Indeed, since we are restricted to $\mathcal{M}(b) = 1$ (see (4.1)) we know that

$$|c_m| < 1. \tag{7.4}$$

Analogously, using the change (4.8), one can see that

$$|p_i| < 2, \ |q_i| < 2 \quad \text{for } i = 1, 2. \tag{7.5}$$

Now, we start by obtaining more accurate upper bounds for each mode.

Lemma 7.5. *Consider the flow* Φ^t *associated to the vector field in* (4.14) *and a point* $(p_1, q_1, q_2, \sigma, c) \in \mathcal{U}_j \subset \Sigma_j^{\mathrm{out}}$*. Then, there exists a constant* $\widetilde{K}_\sigma > 0$ *such that for* $t \in [0, T_{\mathcal{B}_{\mathrm{glob}}^j}]$*,* $\Phi^t(p_1, q_1, \sigma, q_2, c)$ *satisfies*

$$\begin{aligned}
\left|\Phi_{c_k}^t(p_1, q_1, \sigma, q_2, c)\right| &\leq \widetilde{K}_\sigma \widetilde{M}_{\mathrm{ell},\pm}^{(j)} \delta^{(1-r)/2} \quad \textit{for } m \in \mathcal{P}_j^{\pm} \\
\left|\Phi_{c_{j\pm 2}}^t(p_1, q_1, \sigma, q_2, c)\right| &\leq \widetilde{K}_\sigma \widetilde{M}_{\mathrm{adj},\pm}^{(j)} \left(\widetilde{C}^{(j)}\delta\right)^{1/2}
\end{aligned}$$

and

$$\begin{aligned}
\left|\Phi_{p_1}^t(p_1, q_1, \sigma, q_2, c)\right| &\leq \widetilde{K}_\sigma \widetilde{M}_{\mathrm{hyp}}^{(j)} \left(\widetilde{C}^{(j)}\delta\right)^{1/2} \\
\left|\Phi_{q_1}^t(p_1, q_1, \sigma, q_2, c)\right| &\leq \widetilde{K}_\sigma \widetilde{M}_{\mathrm{hyp}}^{(j)} \left(\widetilde{C}^{(j)}\delta\right)^{1/2} \\
\left|\Phi_{p_2}^t(p_1, q_1, \sigma, q_2, c) - p_2^h(t)\right| &\leq \widetilde{K}_\sigma \delta^{r'} \\
\left|\Phi_{q_2}^t(p_1, q_1, \sigma, q_2, c)\right| &\leq \widetilde{K}_\sigma \widetilde{C}^{(j)} \delta \ln(1/\delta).
\end{aligned}$$

We defer the proof of this lemma to the end of the section.

The bounds obtained in Lemma 7.5 are not enough to prove Proposition 7.3 since we need more accurate estimates for the elliptic modes, the future adjacent modes and q_2. We obtain them in the following three lemmas.

Lemma 7.6. *Consider the flow Φ^t associated to the vector field in (4.14) and a point $(p_1, q_1, \sigma, q_2, c) \in \Sigma_j^{\mathrm{out}}$. Then, there exists a constant $\widetilde{K}_\sigma > 0$ such that for $t \in [0, T_{\mathcal{B}^j_{\mathrm{glob}}}]$ and $k \in \mathcal{P}_j^{\pm}$,*

$$\left| \Phi^t_{c_k}(p_1, q_1, \sigma, q_2, c) - c_k e^{iT_{\mathcal{B}^j_{\mathrm{glob}}}} \right| \leq \widetilde{K}_\sigma \delta^{(1-r)/2+r'}.$$

Proof. It is enough to point out that, using the bounds obtained in Lemma 7.5, the equation for c_k in (4.14) can be written as

$$\dot{c}_k = ic_k + \gamma_k(t)$$

where γ satisfies $\|\gamma\|_\infty \leq \widetilde{K}_\sigma \delta^{1-r+r'}$. Then, to finish the proof of the lemma it is enough to apply the variation of constants formula and take into account that the time $T_{\mathcal{B}^j_{\mathrm{glob}}}$ has an upper bound independent of δ. □

Lemma 7.7. *Fix values p_1, q_1, q_2, c_{j-2} and c_k for $k \in \mathcal{P}_j^{\pm}$ such that the set*

$$\mathcal{D} = \{c_1, \ldots, c_{j-2}, p_1, q_1, \sigma, q_2\} \times \widetilde{\mathbb{D}}^{j+2}_{j,-} \times \{c_{j+3}, \ldots, c_{jN}\},$$

where

$$\widetilde{\mathbb{D}}^{j+2}_{j,-} = \left\{ |c_{j+2}| \leq \widetilde{m}^{(j)}_{\mathrm{adj}} \left(\widetilde{C}^{(j)} \delta \right)^{1/2} \right\},$$

satisfies

$$\mathcal{D} \subset \mathcal{U}_j.$$

Consider the flow Φ^t associated to the vector field in (4.14) and define the following map for points in $\mathcal{D}$

$$F_{\mathrm{adj}}(p_1, q_1, \sigma, q_2, c) = \Phi^{T_{\mathcal{B}^j_{\mathrm{glob}}}}_{c_{j+2}}(p_1, q_1, \sigma, q_2, c)$$

Then, there exists $\widetilde{K}_\sigma > 0$ such that

$$\left\{ |c_{j+2}| \leq \widetilde{m}^{(j)}_{\mathrm{adj}} \left(\widetilde{C}^{(j)} \delta \right)^{1/2} / \widetilde{K}_\sigma \right\} \subset F_{\mathrm{adj}}(\mathcal{D}).$$

Proof. Taking into account the estimates obtained in Lemma 7.5, the equation for c_{j+2} in (4.14) can be written as

$$\frac{d}{dt} \begin{pmatrix} c_{j+2} \\ \overline{c_{j+2}} \end{pmatrix} = \begin{pmatrix} ic_{j+2} - i\omega \left(p_2^h(t) \right)^2 \overline{c_{j+2}} + \gamma_{j+2}(t) \\ -i\overline{c_{j+2}} + i\omega^2 \left(p_2^h(t) \right)^2 c_{j+2} + \overline{\gamma}_{j+2}(t) \end{pmatrix},$$

where p_2^h has been defined in (7.3) and γ satisfies $\|\gamma\|_\infty \leq K_\sigma (\widetilde{C}^{(j)} \delta)^{1/2} \delta^{r'}$. Then, to finish the proof it is enough to apply the variation of constants formula. □

Now we obtain the refined estimates for q_2.

Lemma 7.8. *Fix values $p_1, q_1, c_{j\pm 2}$ and c_k for $k \in \mathcal{P}_j^{\pm}$ such that*

$$\mathcal{Q} = \{c_1, \dots, c_{j-2}, p_1, q_1, \sigma\} \times \left[-\widetilde{C}^{(j)}\delta \left(\ln(1/\delta) - \widetilde{m}_{\mathrm{hyp}}^{(j)} \right), -\widetilde{C}^{(j)}\delta \left(\ln(1/\delta) + \widetilde{m}_{\mathrm{hyp}}^{(j)} \right) \right] \times \{c_{j+2}, \dots, c_{j_N}\}$$

satisfies

$$\mathcal{Q} \subset \mathcal{U}_j.$$

Consider the flow Φ^t associated to the vector field in (4.14) and define the following map for points in $\mathcal{Q}$

$$F_{\mathrm{hyp}}(q_2) = \Phi_{q_2}^{T_{\mathcal{B}^j_{\mathrm{glob}}}}(p_1, q_1, \sigma, q_2, c)$$

Then, there exists $\widetilde{K}_\sigma > 0$ and $D^{(j)}$ satisfying

$$\widetilde{C}^{(j)}/\widetilde{K}_\sigma \leq D^{(j)} \leq \widetilde{K}_\sigma \widetilde{C}^{(j)}$$

such that

$$\left[-D^{(j)}\delta \left(\ln(1/\delta) - 1/\widetilde{K}_\sigma \right), -D^{(j)}\delta \left(\ln(1/\delta) + 1/\widetilde{K}_\sigma \right) \right] \subset F_{\mathrm{hyp}}(\mathcal{Q}).$$

Proof. Taking into account the estimates obtained in Lemma 7.5, we write the equation for q_2 in (4.14) as

$$\dot{q}_2 = \zeta_0(t) q_2 + \zeta_1(t),$$

where ζ_0 only depends on p_2^h in (7.3) and ζ_1 satisfies

$$\|\zeta_1\|_\infty \leq \widetilde{K}_\sigma \widetilde{C}^{(j)} \delta.$$

Thus, the proof of the lemma follows from the variation of constants formula. □

We devote the rest of the section to prove Lemma 7.5.

Proof of Lemma 7.5. During the proof of this lemma the time t will always satisfy $t \in [0, T_{\mathcal{B}^j_{\mathrm{glob}}}]$ and the norm $\|\cdot\|_\infty$ will always refer to the supremmum taken over this time interval .

We start by obtaining the bounds for the non-neighbor elliptic modes. By (4.14), one can easily see that for $k \in \mathcal{P}_j^{\pm}$

$$\frac{d}{dt}|c_k|^2 = \frac{1}{2}\left(c_{k-1}^2 + c_{k+1}^2\right)\overline{c_k}^2 - \frac{1}{2}\left(\overline{c_{k-1}}^2 + \overline{c_{k+1}}^2\right) c_k^{\,2}.$$

Then, using (7.4), we have that

$$\frac{d}{dt}|c_k|^2 \leq |c_k|^2$$

and therefore, applying Gronwall estimates we obtain that for $t \in [0, T^j_{\mathrm{glob}}]$,

$$\left|\Phi_{c_k}^t(p_1, q_1, \sigma, q_2, c)\right|^2 \leq e^{T_{\mathcal{B}^j_{\mathrm{glob}}}} |c_k|^2 \leq \widetilde{K}_\sigma \widetilde{M}_{\mathrm{ell},\pm}^{(j)} \delta^{(1-r)}.$$

Proceeding analogously we deal with the adjacent elliptic mode c_{j-2}. Its associated equation is

$$\begin{aligned}\frac{d}{dt}|c_{j-2}|^2 =& \frac{1}{2}c_{j-3}^2\overline{c_{j-2}}^2 + \frac{1}{2}\overline{c_{j-3}}^2 c_{j-2}^{\,2} \\ & - \frac{1}{2}\left(\omega^2 p_1 + \omega q_1\right)^2 \overline{c_{j-2}}^2 - \frac{1}{2}\left(\omega p_1 + \omega^2 q_1\right)^2 c_{j-2}^{\,2}.\end{aligned}$$

Taking into account the bounds in (7.4) and also (7.5), to obtain

$$\frac{d}{dt}|c_{j-2}|^2 \leq 5|c_{j-2}|^2$$

which, applying Gronwall lemma, gives

$$\left|\Phi^t_{c_{j-2}}(p_1, q_1, \sigma, q_2, c)\right|^2 \leq e^{5T_{\mathcal{B}^j_{\rm glob}}}|c_{j-2}|^2 \leq \widetilde{K}_\sigma \widetilde{M}^{(j)}_{\rm adj,-}\widetilde{C}^{(j)}\delta.$$

Analogously, one can obtain

$$\left|\Phi^t_{c_{j+2}}(p_1, q_1, \sigma, q_2, c)\right|^2 \leq e^{5T^j_{\rm glob}}|c_{j+2}|^2 \leq \widetilde{K}_\sigma \widetilde{M}^{(j)}_{\rm adj,+}\widetilde{C}^{(j)}\delta.$$

Now we obtain the bounds for the hyperbolic modes. We define

$$\rho_1(t) = (\Phi^t_{p_1}(p_1, q_1, \sigma, q_2, c), \Phi^t_{q_1}(p_1, q_1, \sigma, q_2, c)).$$

From (4.11), one can see that ρ_1 satisfies an equation of the form $\dot{\rho}_1 = A_1(t)\rho_1$ where $A_1(t)$ is a time dependent matrix (which of course depends on $\Phi^t_{p_1}(p_1, q_1, \sigma, q_2, c)$ itself). Using (7.4) and (7.5), one can deduce that

$$\|A_1\|_\infty \leq \widetilde{K}_\sigma.$$

Then, the fundamental matrix Ψ satisfying $\Psi(0) = \mathrm{Id}$ associated to this system satisfies $\|\Psi\|_\infty \leq \widetilde{K}_\sigma$. Since ρ_1 can be just written

$$\rho_1(t) = \Psi(t)\rho_1(0),$$

using that by hypothesis $|p_1(0)|, |q_1(0)| \leq \widetilde{M}^{(j)}_{\rm hyp}\left(\widetilde{C}^{(j)}\delta\right)^{1/2}$, we have that for $t \in [0, T_{\mathcal{B}^j_{\rm glob}}]$,

$$|\rho_1(t)| \leq \widetilde{K}_\sigma \widetilde{M}^{(j)}_{\rm hyp}\left(\widetilde{C}^{(j)}\delta\right)^{1/2}.$$

We finish the proof of the lemma obtaining the estimates for the (p_2, q_2) components. To this end, let us point out that the equation for q_2 can be written as

$$\dot{q}_2 = a_1(t)q_2 + b_1(t)$$

where $a_1(t)$ and $b_1(t)$ are functions which depend on $\Phi^t_{p_1}(p_1, q_1, \sigma, q_2, c)$. Using (7.5) and the just obtained bounds for the non-neighbor and adjacent elliptic modes and for (p_1, q_1) components, one can easily see that

$$\|a_1\|_\infty \leq \widetilde{K}_\sigma \quad \text{and} \quad \|b_1\|_\infty \leq \widetilde{K}_\sigma\left(\widetilde{C}^{(j)}\delta\right)^{1/2}.$$

Therefore, applying Gronwall lemma, we can deduce that

$$\left|\Phi^t_{q_2}(p_1, q_1, \sigma, q_2, c)\right| \leq \widetilde{K}_\sigma \widetilde{C}^{(j)}\delta \ln(1/\delta).$$

To obtain the bounds for p_2 we define $\xi = p_2 - p_2^h$, where p_2^h is the function defined in (7.3). Using (7.5) and (7.3) we have the a priori bound $\|\xi\|_\infty \leq 3$. Therefore, from (4.14) we can deduce an equation for ξ of the form

$$\dot{\xi} = a_2(t)\xi + b_2(t),$$

where the functions a_2 and b_2 satisfy

$$\|a_2\|_\infty \leq K_\sigma \quad \text{and} \quad \|b_2\|_\infty \leq \widetilde{K}_\sigma \delta^{r'}.$$

Then, applying Gronwall's lemma, we obtain

$$\|\xi\|_\infty \leq \widetilde{K}_\sigma \delta^{r'}$$

which implies the estimate for $\Phi^t_{p_2}(p_1, q_1, \sigma, q_2, c) - p_2^h$. This finishes the proof of the lemma. $\square$

A Proof of Normal Form Theorem 2

In the proof of Theorem 2, we use a generic constant C which depends on η. We consider as a change of variables Γ the time one map of a Hamiltonian vector field X_F, where F is the Hamiltonian

$$F = \frac{1}{4} \sum_{n_1,n_2,n_3,n_4 \in \mathbb{Z}^2} F_{n_1 n_2 n_3 n_4} \alpha_{n_1} \overline{\alpha_{n_2}}, \alpha_{n_3} \overline{\alpha_{n_4}}$$

with coefficients

$$\begin{aligned} F_{n_1n_2n_3n_4} =& \frac{-i}{|n_1|^2 - |n_2|^2 + |n_3|^2 - |n_4|^2} && \text{if } n_1 - n_2 + n_3 - n_4 = 0, \\ & && |n_1|^2 - |n_2|^2 + |n_3|^2 - |n_4|^2 \neq 0 \\ F_{n_1n_2n_3n_4} =& 0 && \text{otherwise.} \end{aligned}$$

The vector field X_F is an analytic vector field from ℓ^1 to iself, which is of order 3 at the origin. Indeed, the a_n component of X_F is given by

$$(X_F)_{\alpha_n} = 2i\partial_{\overline{\alpha}_n} F = 4i \sum_{\substack{n_1,n_2,n_3 \in \mathbb{Z}^2 \\ n_1 - n_2 + n_3 - n = 0 \\ |n_1|^2 - |n_2|^2 + |n_3|^2 - |n_4|^2 \neq 0}} F_{n_1n_2n_3n} \alpha_{n_1} \overline{\alpha_{n_2}} \alpha_{n_3},$$

Then, since $|F_{n_1n_2n_3n}| \leq 1$, we can bound the ℓ^1 norm of X_F as

$$\begin{aligned} \|X_F\|_{\ell^1} &\leq \sum_{n \in \mathbb{Z}^2} |(X_F)_{\alpha_n}| \\ &\leq 4 \sum_{n \in \mathbb{Z}^2} \sum_{\substack{n_1,n_2,n_3 \in \mathbb{Z}^2 \\ n_1 - n_2 + n_3 - n = 0 \\ |n_1|^2 - |n_2|^2 + |n_3|^2 - |n|^2 \neq 0}} |\alpha_{n_1}||\alpha_{n_2}||\alpha_{n_3}| \\ &\leq 4 \sum_{n \in \mathbb{Z}^2} \sum_{\substack{n_1,n_2,n_3 \in \mathbb{Z}^2 \\ n_1 - n_2 + n_3 - n = 0}} |\alpha_{n_1}||\alpha_{n_2}||\alpha_{n_3}| \end{aligned}$$

This last sum is a convolution product of three terms, and therefore, by (3.3), we can conclude that

$$\|X_F\|_{\ell^1} \leq 4\|\alpha\|^3_{\ell^1}.$$

Since $X_F : \ell^1 \to \ell^1$ is an analytic vector field which is small in a neighborhood of the origin, the associated flow Φ^t_F sends the ball $B(\eta)$ to $B(2\eta)$ for $t \in [0, 1]$ and $\eta > 0$ small enough. In particular the change of variables $\Gamma : B(\eta) \to B(2\eta)$ is well defined.

Applying the change Γ to the Hamiltonian H we obtain

$$
\begin{aligned}
\mathcal{H}\circ\Gamma = & H\circ \Phi_F^t\big|_{t=1}\\
=&\mathcal{H}+\{\mathcal{H},F\}+\int_0^1 (1-t)\left\{\{\mathcal{H},F\},F\right\}\circ\Phi_F^t dt\\
=&\mathcal{D}+\mathcal{G}+\{\mathcal{D},F\}\\
&+\{\mathcal{G},F\}+\int_0^1 (1-t)\left\{\{\mathcal{H},F\},F\right\}\circ\Phi_F^t dt,
\end{aligned}
$$

where $\{\cdot,\cdot\}$ denotes the Poisson bracket with respect to the symplectic form $\Omega=\frac{i}{2}\sum_{n\in\mathbb{Z}^2}\alpha_n\wedge\overline{\alpha_n}$. We define

$$
\mathcal{R}=\{\mathcal{G},F\}+\int_0^1 (1-t)\left\{\{\mathcal{H},F\},F\right\}\circ\Phi_F^t dt.
$$

Then, it only remains to obtain the desired bounds for $X_\mathcal{R}$ and Γ and to see that

$$
\mathcal{G}+\{\mathcal{D},F\}=\widetilde{\mathcal{G}}.
$$

To obtain, this last equality, it is enough to use the definition for F to see that

$$
\begin{aligned}
\mathcal{G}+\{\mathcal{D},F\}=&\frac{1}{4}\sum_{n_1-n_2+n_3=n_4}\left(1-i(|n_1|^2-|n_2|^2+|n_3|^2-|n_4|^2)F_{n_1n_2n_3n_4}\right)\alpha_{n_1}\overline{\alpha_{n_2}}\alpha_{n_3}\overline{\alpha_{n_4}}\\
=&\frac{1}{4}\sum_{\substack{n_1-n_2+n_3=n_4\\ |n_1|^2-|n_2|^2+|n_3|^2=|n_4|^2}}\alpha_{n_1}\overline{\alpha_{n_2}}\alpha_{n_3}\overline{\alpha_{n_4}}\\
=&\widetilde{\mathcal{G}}.
\end{aligned}
$$

Now we obtain the bounds for $X_\mathcal{R}$. We start by bounding $X_{\{\mathcal{G},F\}}$, the vector field associated to the Hamiltonian $\{\mathcal{G},F\}$. We have to bound

$$
\left\|X_{\{\mathcal{G},F\}}\right\|_{\ell^1}=2\sum_{n\in\mathbb{Z}^2}\left|\partial_{\overline{\alpha_n}}\{\mathcal{G},F\}\right|.
$$

Then,

$$
\begin{aligned}
\left\|X_{\{\mathcal{G},F\}}\right\|_{\ell^1}\leq&2\sum_{n,m\in\mathbb{Z}^2}\left|\partial_{\overline{\alpha_n}}\left(\partial_{\overline{\alpha_m}}\mathcal{G}\partial_{\alpha_m}F\right)\right|+2\sum_{n,m\in\mathbb{Z}^2}\left|\partial_{\overline{\alpha_n}}\left(\partial_{\alpha_m}\mathcal{G}\partial_{\overline{\alpha_m}}F\right)\right|\\
\leq&2\sum_{n,m\in\mathbb{Z}^2}\left|\partial_{\overline{\alpha_n\alpha_m}}\mathcal{G}\right|\left|\partial_{\alpha_m}F\right|+2\sum_{n,m\in\mathbb{Z}^2}\left|\partial_{\overline{\alpha_m}}\mathcal{G}\right|\left|\partial_{\overline{\alpha_n}\alpha_m}F\right|\\
&+2\sum_{n,m\in\mathbb{Z}^2}\left|\partial_{\overline{\alpha_n}\alpha_m}\mathcal{G}\right|\left|\partial_{\overline{\alpha_m}}F\right|+2\sum_{n,m\in\mathbb{Z}^2}\left|\partial_{\alpha_m}\mathcal{G}\right|\left|\partial_{\overline{\alpha_n\alpha_m}}F\right|.
\end{aligned}
$$

All the terms can be bounded analogously. As an example, we bound the first one,

$$
\begin{aligned}
\sum_{n,m\in\mathbb{Z}^2}\left|\partial_{\overline{\alpha_n\alpha_m}}\mathcal{G}\right|\left|\partial_{\alpha_m}F\right|\leq&4\sum_{n,m\in\mathbb{Z}^2}\left|\sum_{n_1+n_2=m+n}\alpha_{n_1}\alpha_{n_2}\right|\left|\sum_{n_1-n_2+n_3=m}\overline{\alpha_{n_1}}\alpha_{n_2}\overline{\alpha_{n_3}}\right|\\
\leq&4\sum_{n\in\mathbb{Z}^2}\sum_{n_1+n_2=n}|\alpha_{n_1}||\alpha_{n_2}|\sum_{m\in\mathbb{Z}^2}\sum_{n_1-n_2+n_3=m}|\alpha_{n_1}||\alpha_{n_2}||\alpha_{n_3}|,
\end{aligned}
$$

where, in the first line we have taken into account that $|F_{n_1n_2n_3n_4}| \leq 1$. Then, since each sum in the last line is a convolution product, we have that

$$\sum_{n,m\in\mathbb{Z}^2} |\partial_{\overline{\alpha_n\alpha_m}}\mathcal{G}|\,|\partial_{\alpha_m}F| \leq C\|\alpha\|_{\ell^1}^5.$$

Now we bound the other term in $X_{\mathcal{R}}$, which is the vector field $X_{\widehat{\mathcal{R}}}$ associated to

$$\widehat{\mathcal{R}} = \int_0^1 (1-t)\left\{\{\mathcal{H}, F\}, F\right\} \circ \Phi_F^t dt$$

Using that $\{\mathcal{D}, F\} = \widetilde{\mathcal{G}} - \mathcal{G}$, one can write this term as a sum

$$\widehat{\mathcal{R}} = \widehat{\mathcal{R}}_1 + \widehat{\mathcal{R}}_2 + \widehat{\mathcal{R}}_3$$

with

$$\begin{aligned}
\widehat{\mathcal{R}}_1 &= \int_0^1 (1-t)\left\{\widetilde{\mathcal{G}}, F\right\} \circ \Phi_F^t dt\\
\widehat{\mathcal{R}}_2 &= -\int_0^1 (1-t)\left\{\mathcal{G}, F\right\} \circ \Phi_F^t dt\\
\widehat{\mathcal{R}}_3 &= \int_0^1 (1-t)\left\{\{\mathcal{G}, F\}, F\right\} \circ \Phi_F^t dt.
\end{aligned}$$

To bound them, we first obtain bounds for Φ_F^t. The flow satisfies that

$$\Phi_F^t = \mathrm{Id} + \int_0^1 X_F \circ \Phi_F^\tau d\tau.$$

Then, recalling that $\|X_F\|_{\ell^1} \leq 4\|\alpha\|_{\ell^1}^3$, one can easily deduce that

$$\sup_{t\in[0,1]} \|\Phi_F^t - \mathrm{Id}\|_{\ell^1} \leq C\|\alpha\|_{\ell^1}^3.$$

In particular, taking $t=1$, we get the desired estimate for $\Gamma = \Phi_F^1$,

$$\|\Gamma - \mathrm{Id}\|_{\ell^1} \leq C\|\alpha\|_{\ell^1}^3.$$

Finally, to obtain the bounds for the ℓ^1 norms of $X_{\widehat{\mathcal{R}}_j}$, it is enough to write them as convolution products, as done for $X_{\{\mathcal{G},F\}}$, and use the estimate for Φ_F^t. Then, one obtains

$$\begin{aligned}
\left\|X_{\widehat{\mathcal{R}}_1}\right\|_{\ell^1} \leq& C\|\alpha\|^5\\
\left\|X_{\widehat{\mathcal{R}}_2}\right\|_{\ell^1} \leq& C\|\alpha\|^5\\
\left\|X_{\widehat{\mathcal{R}}_3}\right\|_{\ell^1} \leq& C\|\alpha\|^7.
\end{aligned}$$

Thus, we can conclude that

$$\|X_{\mathcal{R}}\|_{\ell^1} \leq C\|\alpha\|^5.$$

This completes the proof.

B Proof of Approximation Theorem 4

We devote this section to proof the Approximation Theorem 4. Even if this proof relies in Gronwall-like estimates as the Approximation result in [CKS+10] (see Lemma 2.3) it presents significant differences. To prove Theorem 4, we need that for large enough time, most of the mass remains supported in the modes in Λ. Namely, that the spreading of mass to other modes is slow enough so that we can still keep track of the growth of Sobolev norms. To achieve this control, as already mentioned in Section 2.4, we strongly take advantage of two facts:

- The condition 6_Λ imposed to the set Λ in Proposition 3.1
- The precise knowledge we have on β^λ in (3.17) thanks to Theorem 3-bis.

Condition 6_Λ avoids that the spreading of mass does not concentrate in some particular modes out of Λ. This could be very harmful because such mode could alter considerably the Sobolev norm. On the other hand, thanks to Theorem 3-bis we know that each β_n^λ with $n \in \Lambda$ it is not small for a *short period of time* (of order $\mathcal{O}(N)$) when the corresponding b_j is a hyperbolic mode (see Section 4). For the rest of the time, which is of order $\mathcal{O}(N^2)$, β_n^λ is considerably smaller and therefore it cannot spread mass to other modes. These improvements allow us to choose the best possible λ to acheive polynomial growth of Sobolev norms.

Now we proceed to prove Theorem 4. Throughout this section C denotes any positive constant independent of N and λ. The solution β^λ is expressed in rotating coordinates (see change (3.7)) and α is not. To compare them in a simpler way, we consider the equation (3.6) in rotating coordinates. To this end, we use that equation (3.4) also preserves the ℓ^2 norm and therefore we perform the change of coordinates

$$\alpha_n = g_n e^{i\left(G+|n|^2\right)t}, \tag{B.1}$$

with $G = -2\|\alpha\|_{\ell^2}^2$. Then, the equation for $g = \{g_n\}_{n\in\mathbb{Z}^2}$ reads

$$-i\dot{g}_n = \mathcal{E}_n(g) + \mathcal{J}_n(g), \tag{B.2}$$

where $\mathcal{E} : \ell^1 \to \ell^1$ is the function defined as

$$\mathcal{E}_n(g) = -|g_n|^2 g_n + \sum_{(n_1,n_2,n_3)\in\mathcal{A}(n)} g_{n_1}\overline{g_{n_2}}g_{n_3} \tag{B.3}$$

with $\mathcal{A}(n) \subset (\mathbb{Z}^2)^3$ defined in (3.8), and $\mathcal{J} : \ell^1 \to \ell^1$ is the vector field associated to the Hamiltonian

$$\mathcal{R}'(g) = \mathcal{R}\left(\left\{g_n e^{i\left(G+|n|^2\right)t}\right\}_{n\in\mathbb{Z}^2}\right),$$

where $\mathcal{R}$ is the Hamiltonian introduced in Theorem 2. Therefore, $\mathcal{J}$ satisfies

$$\|\mathcal{J}(g)\|_{\ell^1} = \mathcal{O}\left(\|g\|_{\ell^1}^5\right). \tag{B.4}$$

Note that equation (B.2) and equation (3.8) only differ by $\mathcal{J}$, that is, in the fifth degree terms of the equation. Moreover, note that $g(0) = \alpha(0)$ and therefore, by the hypotheses of Theorem 4,

$$g(0) = \beta^\lambda(0). \tag{B.5}$$

To prove that g and β are close we define the function ξ as

$$\xi_n = g_n - \beta_n \tag{B.6}$$

and we apply refined Gronwall-like estimates to bound its ℓ^1 norm. Thanks to (B.5), we have that $\xi(0) = 0$. Moreover, from equations (3.8) and (B.2), one can deduce the equation for ξ. It can be written as

$$\dot{\xi} = \mathcal{Z}^0(t) + \mathcal{Z}^1(t)\xi + \mathcal{Z}^2(\xi, t) \tag{B.7}$$

where

$$\mathcal{Z}^0(t) = \mathcal{J}\left(\beta^\lambda\right) \tag{B.8}$$

$$\mathcal{Z}^1(t) = D\mathcal{E}\left(\beta^\lambda\right) \tag{B.9}$$

$$\mathcal{Z}^2(t) = \mathcal{E}\left(\beta^\lambda + \xi\right) - \mathcal{E}\left(\beta^\lambda\right) - D\mathcal{E}\left(\beta^\lambda\right)\xi + \mathcal{J}\left(\beta^\lambda + \xi\right) - \mathcal{J}\left(\beta^\lambda\right). \tag{B.10}$$

Applying the ℓ^1 norm to equation (B.7), we obtain

$$\frac{d}{dt}\|\xi\|_{\ell^1} \leq \left\|\mathcal{Z}^0(t)\right\|_{\ell^1} + \left\|\mathcal{Z}^1(t)\xi\right\|_{\ell^1} + \left\|\mathcal{Z}^2(\xi, t)\right\|_{\ell^1}. \tag{B.11}$$

The next three lemmas give estimates for each term in the right hand side of this equation. Their proofs are deferred to the end of this appendix.

Lemma B.1. *The function $\mathcal{Z}^0$ defined in* (B.8) *satisfies* $\left\|\mathcal{Z}^0\right\|_{\ell^1} \leq C\lambda^{-5}2^{5N}$.

Lemma B.2. *The linear operator $\mathcal{Z}^1(t)$ satisfies $\left\|\mathcal{Z}^1(t)\xi\right\|_{\ell^1} \leq \sum_{n\in\mathbb{Z}^2} f_n(t)|\xi_n|$, where $f_n(t)$ are positive functions satisfying*

$$\int_0^T f_n(t)dt \leq C\gamma N, \tag{B.12}$$

where T is the time given in (3.16) *and γ is the constant given in Theorem 3.*

To obtain estimates for $\mathcal{Z}^2(\xi, t)$ defined in (B.10), we apply bootstrap.

Assume that for $0 < t < T^*$ we have

$$\|\xi(t)\|_{\ell^1} \leq C\lambda^{-3/2}2^{-N}. \tag{B.13}$$

A posteriori we will show that the time (3.16) satisfies $0 < T < T^*$ and therefore the bootstrap assumption holds.

Lemma B.3. *Assume that condition* (B.13) *is satisfied. Then the operator $\mathcal{Z}^2(\xi, t)$ satisfies*

$$\left\|\mathcal{Z}^2(\xi, t)\right\|_{\ell^1} \leq C\lambda^{-5/2}\|\xi(t)\|_{\ell^1}.$$

Combining Lemmas B.1, B.2, B.3, equation (B.11) implies

$$\frac{d}{dt}\|\xi\|_{\ell^1} \leq \sum_{n\in\mathbb{Z}^2}\left(f_n(t) + C\lambda^{-5/2}\right)|\xi_n| + C\lambda^{-5}2^{5N}$$

To obtain bounds for $\|\xi\|_{\ell^1}$ we write this equation as

$$\sum_{n\in\mathbb{Z}^2}\frac{d}{dt}|\xi_n| \leq \sum_{n\in\mathbb{Z}^2}\left(f_n(t) + C\lambda^{-5/2}\right)|\xi_n| + C\lambda^{-5}2^{5N}$$

and we apply a Gronwall-like argument for each harmonic of ξ. Namely, we consider the following change of coordinates,

$$\xi_n = \zeta_n e^{\int_0^t \left(f_n(s)+C\lambda^{-5/2}\right)ds}. \tag{B.14}$$

Then, we obtain

$$\sum_{n\in\mathbb{Z}^2} e^{\int_0^t \left(f_n(s)+C\lambda^{-5/2}\right)ds} \frac{d}{dt}|\zeta_n| \leq C\lambda^{-5}2^{5N}$$

From this equation and taking into account that

$$f_n(t) + C\lambda^{-5/2} \geq 0,$$

we obtain that

$$\frac{d}{dt}\|\zeta\|_{\ell^1} = \sum_{n\in\mathbb{Z}^2} \frac{d}{dt}|\zeta_n| \leq C\lambda^{-5}2^{5N}.$$

Therefore, integrating this equation, taking into account that $\zeta(0) = \xi(0) = 0$ and using the bound for T in (3.16) we obtain that

$$\|\zeta\|_{\ell^1} \leq C\lambda^{-3}2^{5N}\gamma N^2$$

To deduce from this bound, the corresponding bound for $\|\xi\|_{\ell^1}$ it is enough to use the change (B.14), the estimate (B.12) from B.2 and the definition of T in (3.16). Then, we obtain

$$|\xi_n| \leq e^{C\gamma N} e^{\lambda^{-5/2}T}|\zeta_n| \leq 2e^{C\gamma N}|\zeta_n|$$

which implies

$$\|\xi\|_{\ell^1} \leq 2e^{C\gamma N}\|\zeta\|_{\ell^1} \leq 2e^{C\gamma N}\lambda^{-3}2^{5N}\gamma N^2.$$

Therefore, using the condition on λ from Theorem 4 with any $\kappa > C$ and taking N big enough, we obtain that for $t \in [0, T]$

$$\|\xi\|_{\ell^1} \leq \lambda^{-2}$$

and therefore we can drop the bootstrap assumption (B.13).

Finally, taking into account (B.6) and (B.1) we obtain

$$\sum_{n\in\mathbb{Z}^2} \left|\alpha_n e^{-i\left(G+|n|^2\right)t} - \beta_n\right| \leq C\lambda^{-3/2},$$

which is equivalent to statement (3.19) in Theorem 4.

It only remains to prove Lemmas B.1, B.2 and B.3.

Proof of Lemma B.1. Taking into account (B.4), we have that

$$\left\|\mathcal{Z}^0\right\|_{\ell^1} \leq C\left\|\beta^\lambda\right\|_{\ell^1}^5.$$

Therefore it only remains to obtain an upper bound for $\left\|\beta^\lambda\right\|_{\ell^1}$. Taking into account that $\operatorname{supp}\{\beta^\lambda\} \subset \Lambda$, the definition of β^λ in (3.17) and Theorem 3, we have that

$$\left\|\beta^\lambda(t)\right\|_{\ell^1} \leq \sum_{n\in\Lambda} \left|\beta_n^\lambda(t)\right| \leq 2^N\lambda^{-1}\sum_{j=1}^N \left|b_j\left(\lambda^{-2}t\right)\right|.$$

Now it only remains to point out that from the results obtained in Theorem 3–bis, we know that at each time all but three components of b are of size $|b_j| \lesssim \delta^\nu$ for certain $\nu > 0$ whereas the other two satisfy $|b_j| \leq 1$. Then, using the definition of δ in Theorem 3, we obtain that

$$\sum_{j=1}^{N} \left|b_j\left(\lambda^{-2}t\right)\right| \leq C(1+N\delta^\nu) \leq C,$$

which implies

$$\left\|\beta^\lambda(t)\right\|_{\ell^1} \leq C2^N\lambda^{-1}. \tag{B.15}$$

This finishes the proof of the lemma. □

Proof of Lemma B.2. To proof Lemma B.2 we start by analyzing each component of $\mathcal{Z}^1(t)\xi$. To this end, we use the function $\mathcal{E}$ defined in (B.3) to obtain

$$\left(\mathcal{Z}^1(t)\xi\right)_n = \sum_{k\in\mathbb{Z}^2} \partial_{\xi_k}\mathcal{E}_n\left(\beta^\lambda\right)\xi_k + \sum_{k\in\mathbb{Z}^2} \partial_{\overline{\xi}_k}\mathcal{E}_n\left(\beta^\lambda\right)\overline{\xi}_k.$$

We define the functions f_n as

$$f_n(t) = \sum_{k\in\mathbb{Z}^2} \left|\partial_{\xi_n}\mathcal{E}_k\left(\beta^\lambda\right)\right| + \sum_{k\in\mathbb{Z}^2} \left|\partial_{\overline{\xi}_k}\mathcal{E}_n\left(\beta^\lambda\right)\right|. \tag{B.16}$$

We analyze them differently whether $n \in \Lambda$ or $n \notin \Lambda$. We start with the first case.

We fix $n \in \Lambda$ and we want to study which terms in the right hand side of (B.16) are non zero. Indeed, each of the terms $\left|\partial_{\xi_n}\mathcal{E}_k\left(\beta^\lambda\right)\right|$ is of the form $\beta^\lambda_{n_1}\beta^\lambda_{n_2}$ with $(n_1, n_2, n) \in \mathcal{A}(k)$, $(n, n_2, n_1) \in \mathcal{A}(k)$ or $n_1 = n_2 = n = k$ (the last case arising due to the term $-|g_n|^2 g_n$ in (B.3)). Then, these terms are non-zero provided $\beta^\lambda_{n_1} \neq 0$ and $\beta^\lambda_{n_2} \neq 0$. This condition is satisfied provided $n_1, n_2 \in \Lambda$ (see (3.17)). Thus, we have that $n, n_1, n_2 \in \Lambda$. Then, property 1_Λ of the set Λ guarantees that $k \in \Lambda$. Properties 2_Λ and 3_Λ imply that n only belongs to two nuclear families. Therefore, it only interacts with seven vertices (recall that it can interact with itself through the term $-|g_n|^2 g_n$ in (B.3)). This implies that for a fixed n,

$$\partial_{\xi_n}\mathcal{E}_k\left(\beta^\lambda\right) = 0$$

except for seven values of k, which correspond to the parents, children, spouse and sibling of n and n itself. Moreover, for the same reason, each term $\partial_{\xi_n}\mathcal{E}_k\left(\beta^\lambda\right)$ which is non-zero, only contains a finite and independent of N and n number of summands of the form $\beta_{n_1}\overline{\beta}_{n_2}$ with $(n_1, n_2, n) \in \mathcal{A}(k)$, $(n, n_2, n_1) \in \mathcal{A}(k)$ or $n_1 = n_2 = n = k$.

Reasoning in the same way, we can obtain analogous results for the terms $|\partial_{\overline{\xi}_k}\mathcal{E}_n(\beta^\lambda)|$.

From these facts, we can deduce formula (B.12) for $n \in \Lambda$. Indeed, we have seen that f_n only involves seven harmonics of β^λ and that it is quadratic in them. Then, recalling the definition of β^λ in (3.17), Theorem 3-bis ensures that $f_n(t)$ has size $f_n \sim \lambda^{-2}$ for a time interval of order $\lambda^2\ln(1/\delta) \sim \lambda^2\gamma N$ (recall that $\delta = e^{-\gamma N}$) and has size $f_n \sim \lambda^{-2}\delta^\nu \sim \lambda^{-2}e^{-\gamma\nu N}$ for the rest of the time, that is, for a time interval of order $\lambda^2 N\ln(1/\delta) \sim \lambda^2\gamma N^2$. Therefore,

$$\int_0^T f_n(t)dt \leq C\left(N + N^2e^{-\gamma\nu N}\right) \leq C\gamma N.$$

This finishes the proof for $n \in \Lambda$.

Now we need analgous results for $n \notin \Lambda$. We need to see which terms of $\left|\partial_{\xi_n} \mathcal{E}_k\left(\beta^\lambda\right)\right|$, that are of the form $\beta^\lambda_{n_1}\beta^\lambda_{n_2}$, are non-zero. We know that they are non-zero provided $(n_1, n_2, n) \in \mathcal{A}(k)$ or $(n, n_2, n_1) \in \mathcal{A}(k)$ and $n_1, n_2 \in \Lambda$. Note that know the case $n_1 = n_2 = n = k$ is excluded since $n \notin \Lambda$ and $n_1, n_2 \in \Lambda$. Since $n \notin \Lambda$ and $n_1, n_2 \in \Lambda$, property 1_Λ implies that $k \notin \Lambda$. Then, property 6_Λ guarantees that there are at most two rectangles with two vertices in Λ and two out of Λ. Therefore, we have that

$$\partial_{\xi_n} \mathcal{E}_k\left(\beta^\lambda\right) = 0$$

except for three values of k, which correspond to n itself and the other vertex not belonging to Λ of each of these two rectangles. Reasoning as before each term $\partial_{\xi_n} \mathcal{E}_k\left(\beta^\lambda\right)$ which is non-zero, only contains a finite and independent of N and n number of summands of the form $\beta_{n_1}\overline{\beta}_{n_2}$ with $n_1, n_2 \in \Lambda$. Then, reasoning as in the previous case, we obtain

$$\int_0^T f_n(t)dt \leq C\gamma N.$$

This finishes the proof of the lemma. $\square$

Proof of Lemma B.3. To prove Lemma B.3, we split $\mathcal{Z}^2$ in (B.10) as $\mathcal{Z}^2 = \mathcal{Z}^2_1 + \mathcal{Z}^2_2$ with

$$\begin{aligned}
\mathcal{Z}^2_1(t) =& \mathcal{E}\left(\beta^\lambda + \xi\right) - \mathcal{E}\left(\beta^\lambda\right) - D\mathcal{E}\left(\beta^\lambda\right)\xi \\
\mathcal{Z}^2_2(t) =& \mathcal{J}\left(\beta^\lambda + \xi\right) - \mathcal{J}\left(\beta^\lambda\right).
\end{aligned}$$

Using the definition of $\mathcal{E}$ in (B.3), it can be easily seen that

$$\left\|\mathcal{Z}^2_1\right\|_{\ell^1} \leq C\left(\|\beta^\lambda\|_{\ell^1}\|\xi\|^2_{\ell^1} + \|\xi\|^3_{\ell^1}\right).$$

Then, using the bound for $\|\beta^\lambda\|_{\ell^1}$ obtained in (B.15) and the bootstrap assumption (B.13), we obtain

$$\left\|\mathcal{Z}^2_1\right\|_{\ell^1} \leq C\lambda^{-5/2}\|\xi\|_{\ell^1}.$$

We proceed analogously for $\mathcal{Z}^2_2$. Indeed, it satisfies

$$\left\|\mathcal{Z}^2_2\right\|_{\ell^1} \leq C\sum_{k=1}^{5}\|\beta^\lambda\|^{5-k}_{\ell^1}\|\xi\|^k_{\ell^1}$$

and applying (B.15) and (B.13) again, we obtain

$$\left\|\mathcal{Z}^2_2\right\|_{\ell^1} \leq C\lambda^{-5/2}\|\xi\|_{\ell^1}.$$

Thus, we can conclude that

$$\left\|\mathcal{Z}^2\right\|_{\ell^1} \leq C\lambda^{-5/2}\|\xi\|_{\ell^1}.$$

$\square$

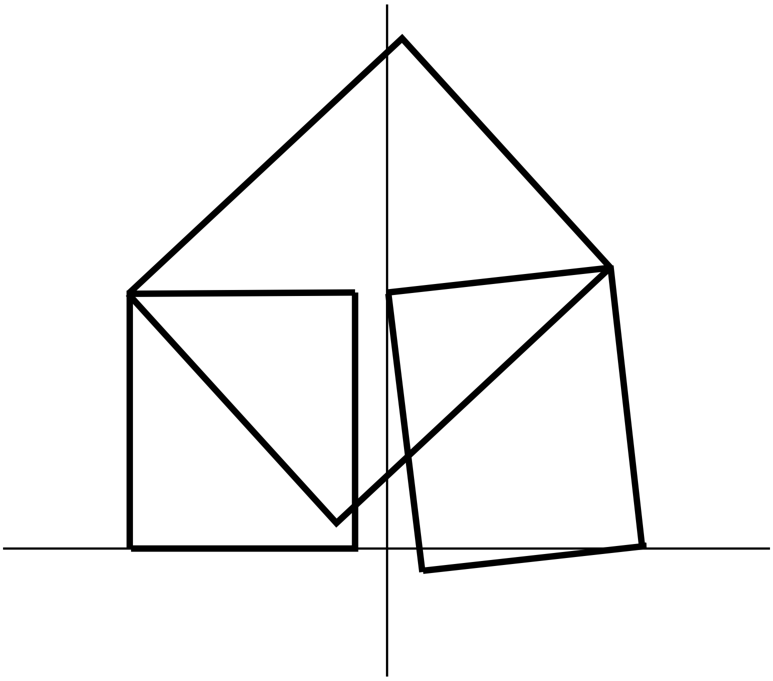

Figure 4: Rectangles

C A result for small initial Sobolev norm

In Theorem 1 we cannot ensure that the initial Sobolev norm $\|u(0)\|_{H^s}$ is arbitrarily small as is done in [CKS+10]. One could impose this condition at the expense of obtaining a worse estimate for the time T. In this appendix we state an analog of Theorem 1 under assuming that $\|u(0)\|_{H^s}$ is arbitrarily small.

Theorem 7. *Let $s > 1$. Then there exists $c > 0$ with the following property: for any small $\mu \ll 1$ and large $\mathcal{A} \gg 1$ there exists a a global solution $u(t,x)$ of (1.1) and a time T satisfying*

$$0 < T \leq \left(\frac{\mathcal{A}}{\mu}\right)^{c \ln(\mathcal{A}/\mu)}$$

such that

$$\|u(T)\|_{H^s} \geq \mathcal{A} \quad \text{and} \quad \|u(0)\|_{H^s} \leq \mu.$$

Remark C.1. *The combination of Theorems 1 and 7 covers all regimes studied in [CKS+10].*

The proof of this theorem follows along the same lines as the proof of Theorem 1 explained in Section 3 taking $\mathcal{K} = \mathcal{A}/\mu$. The only difference is the choice of the parameter λ to ensure

$$\|u(0)\|_{H^s} \leq \mu.$$

Indeed, as it is explained in Section 3, we have that

$$\|u(0)\|_{H^s}^2 \lesssim \lambda^{-2} S_3$$

and, therefore, one needs to choose λ such that $\lambda^{-2} S_3 \sim \mu$. By Proposition 3.1, the constant S_3, defined in (3.20), depends on N. Nevertheless, in that theorem there is no quantitative estimate of this dependence. We will compute it here and show how it affects the estimates for the diffusion time T.

We will show that there is a choice of the set Λ with S_3 from (3.20) satisfying

$$S_3 \lesssim B^{N^2}, \tag{C.1}$$

for certain $B > 0$ independent of N, e.g. $B = 60^4$ applies.

First, using this estimate we derive the time estimate in Theorem 7 from (C.1). Later we prove (C.1). We choose

$$\lambda \sim \frac{1}{\mu} B^{N^2}$$

so that $\lambda^{-2} S_3 \sim \mu$. Then, by Proposition 3.1 we have $N \sim \ln \mathcal{K}$. Taking $\mathcal{K} = \mathcal{A}/\mu$, we know that there exists a constant $c > 0$ such that

$$\lambda \lesssim \left(\frac{\mathcal{A}}{\mu}\right)^{c \ln(\mathcal{A}/\mu)},$$

and therefore, using formula (3.16) we obtain the estimate for the time.

Now we prove (C.1). To this end we use the construction of the set Λ done in [CKS$^+$10]. Recall that the authors first construct the set Λ inside the Gaussian rationals $\mathbb{Q}[i]$ and then multiplying by the least common multiple they map it to the Gaussian integers $\mathbb{Z}[i]$, which is identified with $\mathbb{Z}^2$. Now, we want to place the points in $\mathbb{Q}[i]$ keeping track of the denominators. This gives us the size of the harmonics we are dealing with and, therefore, the size of S_3.

The placement of the modes in $\mathbb{Q}[i]$ is done inductively generation by generation. Namely, we first place Λ_1, then place Λ_2 checking that the conditions $1_\Lambda - 6_\Lambda$ are satisfied, then place Λ_3 and so on. Note that the modes have to be close to the configuration called *prototype embedding* in [CKS$^+$10], Sect. 4, since then we can ensure that (3.9) is satisfied.

First generation: To place the first generation we consider a grid of points in $\mathbb{Q}[i]$ with denominator 60^N. It is clear that we can place Λ_1 in this grid with the points close to the first generation of the *prototype embedding* in [CKS$^+$10]. It can be done so that (co)tangent of a slope between any two points in Λ_1 has numerator and denominator bounded by $Q_1 := 60^N$.

Second generation: The set Λ_1 is divided in pairs of modes which are the parents of different nuclear families. For each of these pairs, we need to place a pair of points of Λ_2 forming a rectangle with the other pair. These new pair is going to be the children of the nuclear family. To place it we consider the circle C having as a diameter the segment between the considered pair in Λ_1. Then, the children have to be placed

- at the endpoints of a different diameter of C.
- they should belong to $\mathbb{Q}[i]$ and
- the conditions $1_\Lambda - 6_\Lambda$ are satisfied.

To see that the children belong to $\mathbb{Q}[i]$, we have to consider a diameter making a Pythagorean angle with the previous diameter, that is an angle θ such that $e^{i\theta} \in \mathbb{Q}[i]$ (see Figure 5).

Let $n = [\sqrt{R/2}]$ be the integer part of $\sqrt{R/2}$. We lower bound the number of θ's whose tangent is rational with numerator and denominator bounded by R as $\sqrt{R/2}$. To see that notice that any triple of the form $a = m^2 - n^2,\ b = 2mn,\ c = m^2 + n^2$ with $m < n$ is Pythagorean. Then there are $n - 1$ values for m giving a Pythagorean triple.

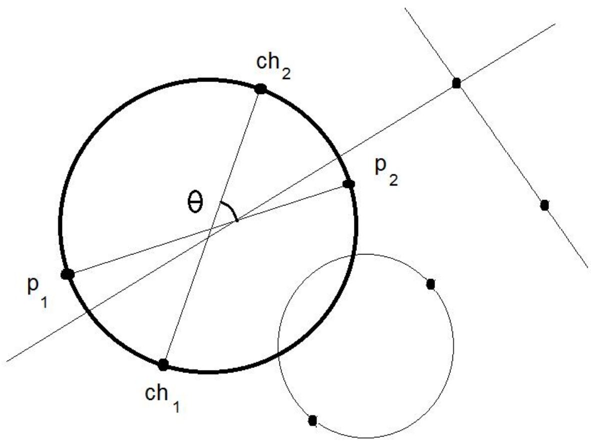

Figure 5: Proper children's choice

Conditions $1_\Lambda - 6_\Lambda$ are satisfied provided the modes in Λ_2 are not placed in certain points of the circle C. The number of these points is of order smaller than 60^N. Indeed, we have to exclude:

- The points of the previous generation (2^N points).
- The points of Λ_2 which have already been placed (at most 2^N).
- To avoid the existence of more rectangles besides the nuclear families, we proceed as follows. We consider
  - all the already placed points,
  - all the lines perpendicular to lines containing two of these points and passing through one of them,
  - all the circles having as a diameter the segment between two of the already placed points (see Figure 5).

  Call $\mathcal{L}$ the set of these lines and $\mathcal{C}$ the set of these circles. The cardinality $|\mathcal{L} \cup \mathcal{C}|$ is at most of order 5^N. Then, we have to exclude all the intersections between any object in $\mathcal{L} \cup \mathcal{C}$ with the circle C.
- To ensure that condition 6_Λ is satisfied, we consider the set $\mathcal{P}$ of the points which are the intersection between any two objects in $\mathcal{L} \cup \mathcal{C}$. It is easy to see that $|\mathcal{P}|$ is of order at most 25^N. Consider the sets
  - $\mathcal{L}'$ containing the lines which are perpendicular to a line containing a point in $\mathcal{P}$ and an already placed point of Λ, and contain one of these two points,
  - $\mathcal{C}'$ containing the circles having as a diameter a segment whose endpoints are a point in $\mathcal{P}$ and an already placed point of Λ.

  The cardinality $|\mathcal{L}' \cup \mathcal{C}'|$ is at most of order 60^N. Then, we have to exclude also the intersections between an element in $\mathcal{L} \cup \mathcal{C}$ and C.

We can place the children of the nuclear family at rational points of the circle C away from the ones just mentioned. To estimate its denominator we apply our estimate on the number of the Pythagorean triples. We have that the number of θ's with slopes whose tangent is given by a rational whose numerator and denominator is bounded by R lower bounded by $\sqrt{R/2}-1$. Thus, we can choose $R = 60^{2N}$. Formula $\tan(\alpha+\beta) = (\tan\alpha + \tan\beta)/(1+\tan\alpha\tan\beta)$ implies that $Q_2 \leq 2\,60^{2N}\,Q_1$. Thus, denominators and numerators in $\Lambda_1 \cup \Lambda_2$ are upper bounded by Q_2. This grid is accurate enough to place the pairs of Λ_2 in the corresponding circles. Iteratively, we can place the following generations refining the grid at each step by dealing with Gaussian rationals whose (co)tangent has numerator and denominator bounded by 60^{3jN} at the j generation. Therefore, after placing the N generations and mapping the set Λ from $\mathbb{Q}[i]$ to $\mathbb{Z}[i]$ we obtain that all the modes $n \in \Lambda$ satisfy

$$|n| \lesssim 60^{3N^2}.$$

This procedure can be done so that the final configuration of modes is close to the *prototype embedding* in [CKS+10] to ensure that condition (3.9) is satisfied. Finally, to obtain the estimate (C.1), it is enough to take any $B \geq 60^4$.

D Notations

- $\mathcal{K}$ — growth of the Sobolev norm of the solution $\|u(t)\|_{H^s}$ from Theorem 1;
- s — index of the Sobolev space.
- $\mathcal{H}$ — the Hamiltonian of (1.1), defined in (3.2);
- $\mathcal{D}$ — quadratic part of the Hamiltonian $\mathcal{H}$ defined in (3.2);
- $\mathcal{G}$ — quartic part of the Hamiltonian $\mathcal{H}$ defined in (3.2);
- $\mathcal{M}$ — abusing notation, mass of both the solutions of the equation (1.1) and of the Toy Model (3.12)
- $\{a_n(t)\}_{n\in\mathbb{Z}^2}$ — Fourier coefficients of the solutions of (1.1) or, equivalently, solution of system $\dot{a}_n = 2\,i\,\partial_{\overline{a_n}}\mathcal{H}$;
- Γ — normal form change for the Hamiltonian (3.2). It is given in Theorem 2.
- $\widetilde{\mathcal{G}}$ — resonant terms of $\mathcal{G}$.
- $\mathcal{R}$ — remainder (of degree 5) of the Hamiltonian $\mathcal{H}$ after performing one step of normal form, that is remainder of the Hamiltonian $\mathcal{H}\circ\Gamma$.
- $\{\alpha_n(t)\}_{n\in\mathbb{Z}^2}$ — Solutions of the normalized Hamiltonian $\mathcal{H}\circ\Gamma$, given in Theorem 2;
- $\mathcal{A}_0(n) \subset (\mathbb{Z}^2)^3$ — collection of the resonance convolutions defined in (3.5);
- $\{\beta_n(t)\}_{n\in\mathbb{Z}^2}$ — rotated fourier coefficients, $\beta_n = \alpha_n e^{-i(G+|n|^2)t}$. They satisfy (3.8).
- $\mathcal{A}(n) \subset (\mathbb{Z}^2)^3$ — collection of reduced resonance convolutions defined after (3.8);
- $N-4$ — number of energy cascades;

- $\Lambda \subset \mathbb{Z}^2$ essential Fourier coefficients given as a disjoint union of N pairwise disjoint generations: $\Lambda = \Lambda_1 \cup \cdots \Lambda_N$. See Proposition 3.1 and preceding discussion.
- $\{b_j(t)\}_{j=1}^N$ solution to the Toy Model (3.12);
- $h(b)$ — Hamiltonian of the Toy Model, given in (3.13);
- $\mathbb{T}_j$ — periodic orbits of the Toy Model (3.12)
- $\{c_k^{(j)}\}_{k\neq j}$ — coordinates adapted to the periodic orbit $\mathbb{T}_j$ after symplectic reduction, given in Section 4.1.
- (p_1, q_1, p_2, q_2) — hyperbolic variables adapted to the periodic orbit $\mathbb{T}_j$ after diagonalization, given in Section 4.1.
- $\mathcal{Z}_{\mathrm{hyp},*}, \mathcal{Z}_{\ell,*}, \mathcal{Z}_{\mathrm{mix},*}$ — types of remainder terms of the original Hamiltonian H after symplectic reduction and diagonalization near the periodic orbit $\mathbb{T}_j$. Subscript means hyperbolic, elliptic and mixed remainder respectively (see Lemma 4.1).
- Σ_j^{in} — transversal section to the stable manifold of $\mathbb{T}_j$, defined in (4.26)
- Σ_j^{out} — transversal section to the unstable manifold of $\mathbb{T}_j$, defined in (4.34)
- $\mathcal{B}^j$ — map from Σ_j^{in} to $\Sigma_{j+1}^{\mathrm{in}}$ given by the flow of the Toy Model (3.12) (see Section 4).
- $\mathcal{B}^j_{\mathrm{loc}}$ — local map from Σ_j^{in} to Σ_j^{out} given by the flow of (3.12), defined in (4.35).
- $\mathcal{B}^j_{\mathrm{glob}}$ — global map from Σ_j^{out} to $\Sigma_{j+1}^{\mathrm{in}}$ given by the flow (3.12), defined in (4.36).
- $a = \mathcal{O}(b)$ means $|b| < Ka$ for some K independent of δ, σ, N, j.
- $a = \mathcal{O}_\sigma(b)$ means $|b| < Ka$ for some K independent of δ, N, j.
- Ψ_{hyp} — the change of coordinates for the hyperbolic Toy Model (see Lemma 5.1).
- Ψ — the change of coordinates for the full Toy Model (see Lemma 6.1).
- $R_{\mathrm{hyp},*}$, $R_{\mathrm{mix},*}$, , $\mathcal{Z}_{\ell,*}$ — collection of remainder terms for the Full Toy Model after normal form transformation Ψ (see Lemma 6.1).
- $\mathcal{V}_j \subset \Sigma_j^{\mathrm{in}}$ — an open subset contained in the domain of definition of $\mathcal{B}^j_{\mathrm{loc}}$ so that $\mathcal{B}^j_{\mathrm{loc}}(\mathcal{V}_j) \subset \mathcal{U}_j$.
- $\mathcal{U}_j \subset \Sigma_j^{\mathrm{out}}$ — an open subset contained in the domain of definition of $\mathcal{B}^j_{\mathrm{glob}}$ so that $\mathcal{B}^j_{\mathrm{glob}}(\mathcal{U}_j) \subset \mathcal{V}_{j+1}$.
- $\mathcal{N}_j^{\pm}$ — initial conditions inside Σ_j^{in} whose orbits under the flow Φ^t have cancellation property (see Lemma 5.2)
- $\mathcal{W}_j$ — an auxiliary set in the (p, q, c)–space (see Corollary 7.2)
- $g_{\mathcal{I}_j}(p_2, q_2, \sigma, \delta)$ — the cancellation function, defined in (6.5) and used in the definition of $\mathcal{N}_j^{\pm}$.

- T_0 — time of evolution of the Toy Model in Theorem 3.
- γ — constant which gives the relation between δ and N.
- $\mathbb{K}$ — constant from upper bound on time in Theorem 3.
- λ — rescaling parameter see (3.15);
- κ — constant wich gives the relation between λ and N.
- T — time of evolution after rescaling, see (3.16);
- $\{b_j^\lambda(t)\}_{j=1}^N$ rescaled solution to the Toy Model, given in (3.15);
- $\{\beta_n^\lambda(t)\}_{n\in\mathbb{Z}^2}$ the lift of the above solution to the Toy Model to approximate solution to (3.8);

Acknowledgements

The authors would like to thank Terence Tao for pointing out an incorrect choice of norms in the original statement of Theorem 1 in an earlier version of the preprint. We thank Jim Colliander for providing several interesting references. A. Simon gave a detailed reading to the paper which improved the exposition.

The authors warmly thank the Fields Insitute and the Institute for Advanced Study for their hospitality, stimulating atmosphere, and support. The first author is partially supported by the Spanish MCyT/FEDER grant MTM2009-06973 and the Catalan SGR grant 2009SGR859. The second author is partially supported by NSF grant DMS-0701271.

References

[Bam97] D. Bambusi. Long time stability of some small amplitude solutions in nonlinear Schrödinger equations. *Comm. Math. Phys.*, 189(1):205–226, 1997.

[Bam99] D. Bambusi. Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations. *Math. Z.*, 230(2):345–387, 1999.

[Bam03] D. Bambusi. Birkhoff normal form for some nonlinear PDEs. *Comm. Math. Phys.*, 234(2):253–285, 2003.

[BB11] M. Berti and L. Biasco. Branching of Cantor manifolds of elliptic tori and applications to PDEs. *Comm. Math. Phys.*, 305(3):741–796, 2011.

[Ber07] M. Berti. *Nonlinear oscillations of Hamiltonian PDEs.* Progress in Nonlinear Differential Equations and their Applications, 74. Birkhäuser Boston Inc., Boston, MA, 2007.

[BG06] D. Bambusi and B. Grébert. Birkhoff normal form for partial differential equations with tame modulus. *Duke Math. J.*, 135(3):507–567, 2006.

[BK92] I. U. Bronsteĭn and A. Ya. Kopanskiĭ. Finitely smooth normal forms of vector fields in the vicinity of a rest point [MR1167014 (93j:58120)]. In *Global analysis—studies and applications, V*, volume 1520 of *Lecture Notes in Math.*, pages 157–172. Springer, Berlin, 1992.

[BK94] I. U. Bronstein and A. Ya. Kopanskiĭ. *Smooth invariant manifolds and normal forms*, volume 7 of *World Scientific Series on Nonlinear Science. Series A: Monographs and Treatises.* World Scientific Publishing Co. Inc., River Edge, NJ, 1994.

[Bou93] J. Bourgain. Fourier transform restriction phenomena for certain lattice subsets and application to nonlinear evolution equations. i. schrödinger equations. *Geom. Funct. Anal.*, 3(2):107–156, 1993.

[Bou96] J. Bourgain. On the growth in time of higher Sobolev norms of smooth solutions of Hamiltonian PDE. *Internat. Math. Res. Notices*, 6:277–304, 1996.

[Bou98] J. Bourgain. Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations. *Ann. of Math. (2)*, 148(2):363–439, 1998.

[Bou00a] J. Bourgain. On diffusion in high-dimensional Hamiltonian systems and PDE. *J. Anal. Math.*, 80:1–35, 2000.

[Bou00b] J. Bourgain. Problems in Hamiltonian PDE's. *Geom. Funct. Anal.*, Special Volume, Part I:32–56, 2000. GAFA 2000 (Tel Aviv, 1999).

[Bou04] J. Bourgain. Remarks on stability and diffusion in high-dimensional Hamiltonian systems and partial differential equations. *Ergodic Theory Dynam. Systems*, 24(5):1331–1357, 2004.

[CDKS01] J. E. Colliander, J.-M. Delort, C. E. Kenig, and G. Staffilani. Bilinear estimates and applications to 2D NLS. *Trans. Amer. Math. Soc.*, 353(8):3307–3325 (electronic), 2001.

[CE10] R. Carles and Faou E. Energy cascade for nls on $\mathbb{T}^d$. page 15 pp. Arxiv: 1010.5173v1, 2010.

[CKO12] J. Colliander, S. Kwon, and T. Oh. A remark on normal forms and the upside-down I-method for periodic NLS: growth of higher Sobolev norms. Preprint available at `http://arxiv.org/abs/1010.2501`, 2012.

[CKS$^+$10] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao. Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation. *Invent. Math.*, 181(1):39–113, 2010.

[CW93] W. Craig and C.E. Wayne. Newton's method and periodic solutions of nonlinear wave equations. *Comm. Pure Appl. Math.*, 46(11):1409–1498, 1993.

[CW10] F. Catoire and W.-M. Wang. Bounds on Sobolev norms for the defocusing nonlinear Schrödinger equation on general flat tori. *Commun. Pure Appl. Anal.*, 9(2):483–491, 2010.

[EK10] L. H. Eliasson and S. B. Kuksin. KAM for the nonlinear Schrödinger equation. *Ann. of Math. (2)*, 172(1):371–435, 2010.

[Fen77] N. Fenichel. Asymptotic stability with rate conditions. II. *Indiana Univ. Math. J.*, 26(1):81–93, 1977.

[Fen74] N. Fenichel. Asymptotic stability with rate conditions. *Indiana Univ. Math. J.*, 23:1109–1137, 1973/74.

[GG10] P. Gérard and S. Grellier. The cubic Szegő equation. *Ann. Sci. Éc. Norm. Supér. (4)*, 43(5):761–810, 2010.

[GIP09] B. Grébert, R. Imekraz, and E. Paturel. Normal forms for semilinear quantum harmonic oscillators. *Comm. Math. Phys.*, 291(3):763–798, 2009.

[GKP12] B. Grébert, T. Kappeler, and J. Pöschel. Normal form theory for the NLS equation. Preprint available at `http://arxiv.org/abs/0907.39381`, 2012.

[Han11] Z. Hani. Global and dynamical aspects of nonlinear schrödinger equations on compact manifolds. *Ph. D. thesis UCLA*, 2011.

[Han12] Z. Hani. Long-time instability and unbounded sobolev orbits for some periodic nonlinear schrdinger equations. Preprint available at `http://arxiv.org/abs/1210.7509`, 2012.

[HPS77] M.W. Hirsch, C.C. Pugh, and M. Shub. *Invariant manifolds*, volume 583 of *Lecture Notes in Math.* Springer-Verlag, Berlin, 1977.

[KLS11] V. Kaloshin, M. Levi, and M. Saprykina. Arnold diffusion in a pendulum lattice. In *preprint*, page 22pp. 2011.

[KP96] S. Kuksin and J. Pöschel. Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation. *Ann. of Math. (2)*, 143(1):149–179, 1996.

[KP03] T. Kappeler and J. Pöschel. *KdV & KAM.* Springer-Verlag, Berlin, 2003.

[Kuk93] S.B. Kuksin. *Nearly integrable infinite-dimensional Hamiltonian systems*, volume 1556 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 1993.

[Kuk95] S. B. Kuksin. On squeezing and flow of energy for nonlinear wave equations. *Geom. Funct. Anal.*, 5(4):668–701, 1995.

[Kuk96] S. Kuksin. Growth and oscillations of solutions of nonlinear Schrödinger equation. *Comm. Math. Phys.*, 178(2):265–280, 1996.

[Kuk97a] S. B. Kuksin. On turbulence in nonlinear Schrödinger equations. *Geom. Funct. Anal.*, 7(4):783–822, 1997.

[Kuk97b] S. B. Kuksin. Oscillations in space-periodic nonlinear Schrödinger equations. *Geom. Funct. Anal.*, 7(2):338–363, 1997.

[Kuk99] S. B. Kuksin. Spectral properties of solutions for nonlinear PDEs in the turbulent regime. *Geom. Funct. Anal.*, 9(1):141–184, 1999.

[Poc11] O. Pocovnicu. Explicit formula for the solution of the Szegö equation on the real line and applications. *Discrete Contin. Dyn. Syst.*, 31(3):607–649, 2011.

[Poc12] O. Pocovnicu. First and second order approximations for a nonlinear wave equation. Preprint available at `http://arxiv.org/abs/1111.6060`, 2012.

[PP12] C. Procesi and M. Procesi. Normal Form for the Schrödinger equation with analytic non–linearities. Preprint available at `http://arxiv.org/abs/1012.0446`, 2012.

[Rab78] P. H. Rabinowitz. Free vibrations for a semilinear wave equation. *Comm. Pure Appl. Math.*, 31(1):31–68, 1978.

[Šil67] L. P. Šil′nikov. On a problem of Poincaré-Birkhoff. *Mat. Sb. (N.S.)*, 74 (116):378–397, 1967.

[Soh10a] V. Sohinger. Bounds on the growth of high Sobolev norms of solutions to 2D Hartree Equations. Preprint available at `http://arxiv.org/abs/1003.5709`, 2010.

[Soh10b] V. Sohinger. Bounds on the growth of high Sobolev norms of solutions to Nonlinear Schrodinger Equations on $\mathbb{R}$. Preprint available at `http://arxiv.org/abs/1003.5707`, 2010.

[Soh11] V. Sohinger. Bounds on the growth of high Sobolev norms of solutions to nonlinear Schrödinger equations on S^1. *Differential Integral Equations*, 24(7-8):653–718, 2011.

[Sta97a] G. Staffilani. On the growth of high sobolev norms of solutions for kdv and schrödinger equations. *Duke Math Journal*, 86(1):109–42, 1997.

[Sta97b] G. Staffilani. Quadratic forms for a 2-D semilinear Schrödinger equation. *Duke Math. J.*, 86(1):79–107, 1997.

[Way90] C.E. Wayne. Periodic and quasi-periodic solutions of nonlinear wave equations via kam theory. *Comm. Math. Phys.*, 127(3):479–528, 1990.

[Zho08] S. Zhong. The growth in time of higher Sobolev norms of solutions to Schrödinger equations on compact Riemannian manifolds. *J. Differential Equations*, 245(2):359–376, 2008.

[ZS71] V. E. Zakharov and A. B. Shabat. Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media. *Ž. Èksper. Teoret. Fiz.*, 61(1):118–134, 1971.